Submitted to the Annals of Applied Probability
arXiv: math.PR/1112.6330

THE DIAMETER OF WEIGHTED RANDOM GRAPHS

BY HAMED AMINI AND MARC LELARGE

EPFL and INRIA

In this paper we study the impact of random exponential edge weights on the distances in a random graph and, in particular, on its diameter. Our main result consists of a precise asymptotic expression for the maximal weight of the shortest weight paths between all vertices (the weighted diameter) of sparse random graphs, when the edge weights are i.i.d. exponential random variables.

1. Introduction and main results. Real-world networks are described not only by their graph structure, which give us information about valid links between vertices in the network, but also by their associated edge weights, representing cost or time required to traverse the edge. The analysis of the asymptotics of typical distances in edge weighted graphs has received much interest by the statistical physics community in the context of *first passage percolation problems*. First-passage percolation (FPP) describes the dynamics of a fluid spreading within a random medium. In this paper we study the impact of random exponential edge weights on the distances in a random graph and, in particular, on its diameter.

The typical distance and diameter of non-weighted graphs have been studied by many people, for various models of random graphs. A few examples are the results of Bollobás and Fernandez de la Vega [12], van der Hofstad, Hooghiemstra and Van Mieghem [19], Fernholz and Ramachandran [16], Chung and Lu [14], Bollobás, Janson and Riordan [13] and Riordan and Wormald [30]. First passage percolation model has been mainly studied on lattices motivated by its subadditive property and its link to a number of other stochastic processes, see e.g., [17, 26, 18] for a more detailed discussion. First passage percolation with exponential weights has received substantial attention (see e.g. [5, 19, 6, 8, 20, 7]), in particular on the complete graph, and, more recently, also on random graphs.

A weighted graph (G, w) is the data of a graph $G = (V, E)$ and a collection of weights $w = \{w_e\}_{e \in E}$ associated to each edge $e \in E$. We suppose that all the edge weights are non-negative. For two vertices a and $b \in V$, a path between a and b is a sequence $\pi = (e_1, e_2, \dots e_k)$ where $e_i = \{v_{i-1}, v_i\} \in E$

AMS 2000 subject classifications: Primary 60C05, 05C80; secondary 90B15
Keywords and phrases: first passage percolation, weighted diameter, random graphs

and $v_i \in V$ for $i \in \{1, \ldots, k\} = [1, k]$, with $v_0 = a$ and $v_k = b$. We write $e \in \pi$ if the edge $e \in E$ belongs to the path π, i.e., if $e = e_i$ for an $i \in [1, k]$. For $a, b \in V$, the weighted distance between a and b is given by

$$\mathrm{dist}_w(a, b) = \mathrm{dist}_w(a, b; G) = \min_{\pi \in \Pi(a,b)} \sum_{e \in \pi} w_e \ ,$$

where the minimum is taken over all the paths between a and b in the graph G. The *weighted diameter* is then given by

$$\mathrm{diam}_w(G) = \max\{\mathrm{dist}_w(a, b),\ a, b \in V,\ \mathrm{dist}_w(a, b) < \infty\} \ ,$$

and the *weighted flooding time* for $a \in V$ is defined by

$$\mathrm{flood}_w(a, G) = \max\{\mathrm{dist}_w(a, b),\ b \in V,\ \mathrm{dist}_w(a, b) < \infty\}.$$

1.1. *Random graphs with given degree sequence.* For $n \in \mathbb{N}$, let $(d_i)_1^n$ be a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even. By means of the configuration model (Bender and Canfield [4], Bollobás[10]), we define a random multigraph with given degree sequence $(d_i)_1^n$, denoted by $G^*(n, (d_i)_1^n)$ as follows: to each node $i \in [1, n]$ we associate d_i labeled half-edges. All half-edges need to be paired to construct the graph, this is done by uniformly matching them. When a half-edge of i is paired with a half-edge of j, we interpret this as an edge between i and j. The graph $G^*(n, (d_i)_1^n)$ obtained following this procedure may not be simple, i.e., may contain self-loops due to the pairing of two half-edges of i, and multi-edges due to the existence of more than one pairing between two given nodes. Conditional on the multigraph $G^*(n, (d_i)_1^n)$ being a simple graph, we obtain a uniformly distributed random graph with the given degree sequence, which we denote by $G(n, (d_i)_1^n)$, [23]. We consider asymptotics as the numbers of vertices tend to infinity, and thus we assume throughout the paper that we are given, for each n, a sequence $\mathbf{d}^{(n)} = (d_i^{(n)})_1^n = (d_i)_1^n$ of non-negative integers such that $\sum_{i=1}^n d_i^{(n)}$ is even. For notational simplicity we will sometimes not show the dependency on n explicitly.

For $k \in \mathbb{N}$, let $u_k^{(n)} = |\{i,\ d_i = k\}|$ be the number of vertices of degree k. From now on, we assume that the sequence $(d_i)_1^n$ satisfies the following regularity conditions analogous to the ones introduced in [29]:

Condition 1.1. For each n, $\mathbf{d}^{(n)} = (d_i^{(n)})_1^n = (d_i)_1^n$ is a sequence of positive integers such that $\sum_{i=1}^n d_i$ is even and, for some probability distribution $(p_r)_{r=1}^\infty$ over integers independent of n and with finite mean $\mu := \sum_{k \geq 1} k p_k \in [1, \infty)$, the following holds:

(i) $u_k^{(n)}/n \to p_k$ for every $k \geq 1$ as $n \to \infty$;
(ii) For some $\epsilon > 0$, $\sum_{i=1}^n d_i^{2+\epsilon} = O(n)$.

Note that the condition $d_i \geq 1$ for all i, is not restrictive since removing all isolated vertices from a graph will not affect the (weighted) distances.

1.2. *Main results.* We define $q = \{q_k\}_{k=0}^\infty$ the size-biased probability mass function corresponding to p, by

$$\forall k \geq 0, \quad q_k := \frac{(k+1)p_{k+1}}{\mu}, \tag{1.1}$$

and let ν denote its mean:

$$\nu := \sum_{k=0}^\infty k q_k \in (0, \infty) \qquad \text{(By Condition 1.1(ii))}. \tag{1.2}$$

Let $\phi_p(z)$ be the probability generating function of $\{p_k\}_{k=0}^\infty$: $\phi_p(z) = \sum_{k=0}^\infty p_k z^k$, and let $\phi_q(z)$ be the probability generating function of $\{q_k\}_{k=0}^\infty$: $\phi_q(z) = \sum_{k=0}^\infty q_k z^k = \phi_p'(z)/\mu$. In this paper, we will consider only the case where $\nu > 1$. In particular, there exists a unique λ in $(0,1)$ such that $\lambda = \phi_q(\lambda)$ and if $\mathcal{C}$ is the size (in number of vertices) of the largest component of $G(n, (d_i)_1^n)$, then we have by [Molloy, Reed [29] - Janson, Luczak [23]]: $\mathcal{C}/n \xrightarrow{p} 1 - \phi_p(\lambda) > 0$. In addition, we introduce

$$\lambda_* = \phi_q'(\lambda) = \sum_{k=1}^\infty k q_k \lambda^{k-1} \in [0,1)\,. \tag{1.3}$$

We can now state our main theorem.

THEOREM 1.2. *Let $(G(n,(d_i)_1^n), w)$ be a sequence of random weighted graphs where $w = \{w_e\}_{e\in E}$ are i.i.d. rate one exponential random variables. Assume* Condition 1.1 *and that ν defined in (1.2) is such that $\nu > 1$. Assume that all the graphs have the same minimum degree denoted by $d_{\min} = \min_{i\in[1,n]} d_i$ and moreover that $p_{d_{\min}} > 0$. Let $\Gamma : \mathbb{N}^* \to \mathbb{R}$ be defined by:*

$$\Gamma(d) := d\,\mathbb{1}[d \geq 3] + 2(1-q_1)\,\mathbb{1}[d=2] + (1-\lambda_*)\,\mathbb{1}[d=1]. \tag{1.4}$$

Let a, b be two uniformly chosen vertices in this graph. If we condition the vertices a and b to be connected, we have

$$\frac{\mathrm{dist}_w(a,b;G(n,(d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\nu - 1}. \tag{1.5}$$

If we condition the vertex a to be in the largest component, we have

$$\frac{\mathrm{flood}_w(a, G(n,(d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\nu-1} + \frac{1}{\Gamma(d_{\min})}. \tag{1.6}$$

Finally, we have:

$$\frac{\mathrm{diam}_w(G(n,(d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\nu-1} + \frac{2}{\Gamma(d_{\min})}. \tag{1.7}$$

REMARK 1.3. Note that $\nu > 1$ implies that $\sum_{k=0}^{\infty} k(k-2)p_k > 0$ so that there is a positive fraction of nodes in $G(n,(d_i)_1^n)$ with degree 3 or larger. In particular, we have $q_1 = 2p_2/\mu < 1$ and $\lambda_* < 1$ so that we have $\Gamma(d) > 0$ for all $d \in \mathbb{N}^* = \{1, 2 \dots\}$.

We now comment our result with respect to related literature. Our main contribution is (1.7) while results (1.5) and (1.6) follow from the analysis required to prove (1.7). Indeed, a much stronger version of (1.5) has been proved for a sligthly different model of random graphs by Bhamidi, van der Hofstad, and Hooghiemstra in [8]. Theorem 3.1 in [8] shows that if the sequence $(d_i)_1^n$ is a sequence of i.i.d. (nondegenerate) random variables with $d_{\min} \geq 2$ and finite variance, then there exists a random variable V such that (conditioning on a and b being connected)

$$\mathrm{dist}_w(a, b; G^*(n,(d_i)_1^n))) - \frac{\log n}{\nu-1} \xrightarrow{d} V.$$

We expect this result to be valid for our model of random graphs $G(n,(d_i)_1^n)$ where the degrees d_i satisfy Condition 1.1 (but we did not try to prove it). [8] and [7] give also results when the degree sequence has no finite second moment and no finite first moment.

Motivated by the analysis of the diameter of the largest component of a critical Erdős-Rényi random graph (without edge weights), Ding, Kim, Lubetzky, and Peres [15] show that if $d_i = r \geq 3$ for all i, then we have with high probability:

$$\mathrm{diam}_w(G^*(n,r)) = \left(\frac{1}{r-2} + \frac{2}{r}\right)\log n + O(\log\log n).$$

The intuition behind this formula is simple: consider a vertex in $G^*(n,r)$, its closest neighbor is at distance given by an exponential random variable with rate r (i.e. the minimun of r exponential rate one random variables). Hence the probability for this distance to be larger than $\log n/r$ is n^{-1}. Since there are n vertices with degree r, a simple argument shows that we will find two nodes with closest neighbors at distance $\log n/r$. The diameter will be obtained by taking a shortest path between these two nodes. Each such node will first give a contribution of $\log n/r$ to reach its closest neighbor and then the path between these neighbors will be typical, of the order $\log n/(r-2)$. This simple heuristic argument shows that our result on the diameter depends crucially on the weights being exponentially distributed or at least have an exponential tail. We refer to [9] for recent results on distances with i.i.d. weights. As we will see, the presence of nodes with degree one and two makes the analysis much more involved than in [15]. As soon as, a fraction of nodes have degree two, there will be long paths constitued by a chain of such nodes and we will see that these paths contribute to the diameter.

In [2], this result is used to analyze an asynchronous randomized broadcast algorithm for random regular graphs. In continuous-time, each node is endowed with a Poisson point process with rate 1 and contacts one of its neighbors uniformly at random at each point of his process. In a push model, if a node holds the message, it passes the message to its randomly chosen neighbor regardless of its state. The results in [2] show that the asynchronous version of the algorithm performs better than its synchronized version: in the large size limit of the graph, it will reach the whole network faster even if the local dynamics are similar on average.

We end this section by a simple remark. Our results can be applied to some other random graphs models too by conditioning on the degree sequence. In particular, our results will apply whenever the random graph conditioned on the degree sequence has a uniform distribution over all possibilities. Notable examples of such graphs are $G(n,p)$, the Bernoulli random graph with n vertices and edge probability p and $G(n,m)$, the uniformly random graph with n vertices and m edges. For example, for $G(n,p)$ with $np \to \mu \in (0,\infty)$ or $G(n,m)$ with $2m/n \to \mu$, the Condition 1.1(i) holds in probability with (p_k) a Poisson distribution with parameter μ, $p_k = e^{-\mu}\frac{\mu^k}{k!}$. In Appendix B, we show that thanks to Skorohod coupling theorem ([25, Theorem 3.30]) our results still apply in this setting. By taking care of removing isolated nodes, our result gives in this case (note that $\phi_q(z) = e^{-\mu(1-z)}$):

THEOREM 1.4. *Let $\mu > 1$ be fixed, and let $\lambda_* < 1$ satisfy $\lambda_* e^{-\lambda_*} = \mu e^{-\mu}$. Assume $G_n = G(n,p)$ where $np \to \mu \in (0,\infty)$ (or $G_n = G(n,m)$ with $2m/n \to \mu \in (0,\infty)$) with i.i.d. rate 1 exponential weights on its edges.*

Then

$$\frac{\mathrm{diam}_w(G_n)}{\log n} \xrightarrow{p} \frac{1}{\mu-1} + \frac{2}{1-\lambda_*}. \tag{1.8}$$

This result improves on a lower bound of the weighted diameter given by Bhamidi, van der Hofstad and Hooghiemstra in [6, Theorem 2.6]. Note that [6] also deals with the case $np \to \infty$ which is out of the scope of the present paper.

1.3. *Overview of the proof and organization of the paper.* Our work is a direct generalization of [15] with significantly more involved calculations. The first key idea of the proof from [15] is to grow balls centered at all vertices of the graph simultaneously. The time when two balls centered at a and b respectively intersect is exactly the half of the weighted distance between a and b. (In what follows, we will sometimes deliberately use the term time instead of the term weighted distance.) Hence the weighted diameter becomes twice the time when the last two balls intersect. A simple argument shows that any two balls containing slightly more than $\sqrt{n}$ vertices ($2\sqrt{rn\log n}$ vertices for r-regular case) will intersect with high probability; see Proposition 3.1. Hence it will be enough to control the time at which all balls have reached this critical size of order $\sqrt{n}$ in order to prove an upper bound for the weighted diameter. For a proof of the upper bound on the diameter, we apply an union bound argument as in [15]. Hence, we need to find the right time such that the probability for a (typical) ball to reach size $\sqrt{n}$ is of order n^{-1}. In order to do so, we use the second main idea of the proof: we couple the exploration process on the weighted graph with a continuous time Markov branching process. This coupling argument is quite standard and we will deal here with the same branching process approximation for the exploration process on the graph as in [8]. However, we are facing here new difficulties as we need to consider events here to of small probability for this exploration process (of order n^{-1}). In particular, we need to show that the coupling is still valid for such large deviations. When $d_{\min} \geq 3$, the argument of [15] can be extended easily [2]. But as soon as $d_{\min} \leq 2$, several complications happen. First as shown in [3], the asymptotics for the large deviations of the branching process depend on the minimal possible offspring. Second, as soon as $d_{\min} = 1$, the small components of the graph contain now a positive fraction of the nodes. We need to bound the diameter of these small components and to study the diameter on the largest component, we need to condition our exploration process on 'non-extinction'. Similarly, the presence of degree one nodes, significantly complicates the proof of the lower bound. In order to apply the second moment method as in [15], we need to first remove vertices with degree one iteratively to work with the 2-core of the graph (indeed an augmented version of this 2-core, see Section 4.2 for details).

We consider in Section 2 the exploration process for configuration model which consists in growing balls simultaneously from each vertex. A precise treatment of the exploration process, resulting in information about the growth rates of the balls are given in this section. In addition, the section provides some necessary notations and definitions that will be used throughout the last three sections. Sections 3 and 4 form the heart of the proof. We first prove that the above bound is an upper bound for the weighted diameter. This will consist in defining the two parameters α_n and β_n with the following significance. (i) Two balls of size at least β_n intersect almost surely, (ii) considering the growing balls centered at a vertex in the graph, the time it takes for the balls to go from size α_n to size β_n have all the same asymptotic for all the vertices of the graph, and the asymptotic is half of the typical weighted distance in the graph, and (iii) the time it takes for the growing balls centered at a given vertex to reach size at least α_n is upper bounded by $\frac{1+\epsilon}{\Gamma(d_{\min})}\log n$ for all $\epsilon > 0$ with high probability (w.h.p.). This will show that the diameter is w.h.p. bounded above by $(1+\epsilon)(\frac{1}{\nu-1} + \frac{2}{\Gamma(d_{\min})})\log n$, for all $\epsilon > 0$. The last section provides the corresponding lower bound. To obtain the lower bound, we show that w.h.p. (iv) there are at least two nodes with degree $d_{\min}$ such that the time it takes for the balls centered at these vertices to achieve size at least α_n is worst than the other vertices, and is lower bounded by $\frac{1-\epsilon}{\Gamma(d_{\min})}\log n$, for all $\epsilon > 0$. And using this, we conclude that the diameter is w.h.p. bounded below by $(1-\epsilon)(\frac{1}{\nu-1} + \frac{2}{\Gamma(d_{\min})})\log n$, for all fixed $\epsilon > 0$, finishing the proof of our main theorem.

The actual values of α_n and β_n will be

$$\alpha_n := \lfloor \log^3 n \rfloor, \text{ and } \beta_n := \lfloor 3\sqrt{\frac{\mu}{\nu-1} n \log n} \rfloor. \tag{1.9}$$

When $d_{\min} = 1$, the longest shortest path in a random graph will be between a pair of vertices a and b of degree one. Furthermore, this path consists of a path from a to the 2-core, a path through the 2-core, and a path from the 2-core to b. For this, we need to provide some preliminary results on the structure of the 2-core, this is done in Appendix A. In Appendix B, we show that our results still apply for random graphs $G(n,p)$ and $G(n,m)$ by conditioning on the degree sequence.

Basic notations. We usually do not make explicit reference to the probability space since it is usually clear to which one we are referring. We say that an event A holds almost surely, and we write a.s., if $\mathbb{P}(A) = 1$. The indicator function of an event A is of particular interest, and it is denoted by $\mathbb{1}[A]$. We consider the asymptotic case when $n \to \infty$ and say that an event holds w.h.p. (with high probability) if it holds with probability tending to 1 as $n \to \infty$. We denote by $\xrightarrow{d}$, and $\xrightarrow{p}$, convergence in distribution, and in probability, respectively. Similarly, we use o_p and O_p in a standard way. For example, if (X_n) is a sequence of random variables, then $X_n = O_p(1)$ means that "X_n is bounded in probability" and $X_n = o_p(n)$ means that $X_n/n \xrightarrow{p} 0$.

2. First Passage Percolation in $G^*(n, (d_i)_1^n)$. We start this section by introducing some new notations and definitions. Before this, one remark is in order. In what follows, we will sometimes deliberately use the term "time" instead of the term "weighted distance". It will be clear from the context what we actually mean by this.

Let $\big(G = (V, E), w\big)$ be a weighted graph. For a vertex $a \in V$ and a real number $t > 0$, the t-radius neighborhood of a in the (weighted) graph, or the ball of radius t centered at a, is defined as

$$B_w(a, t) := \big\{\, b,\ \mathrm{dist}_w(a, b) \le t \,\big\}.$$

The first time t where the ball $B_w(a, t)$ reaches size $k + 1$ will be denoted by $T_a(k)$ for $k \ge 0$, i.e.

$$T_a(k) = \min\big\{\, t : |B_w(a, t)| \ge k + 1 \,\big\}, \qquad T_a(0) = 0.$$

If there is no such t, i.e. if the component containing a has size at most k, we define $T_a(k) = \infty$. More precisely, we use I_a to denote the size of the component containing a in the graph minus one, in other words,

$$I_a := \max\big\{\, |B_w(a, t)|, t \ge 0 \,\big\} - 1,$$

so that for all $k > I_a$, we set $T_a(k) = \infty$. Note that there is a vertex in $B_w(a, T_a(k))$ which is not in any ball of smaller radius around a. When the weights are i.i.d. according to a random variable with continuous density, this vertex is in addition unique with probability one. We will assume this in what follows. For an integer $i \le I_a$, we use $\widehat{d}_a(i)$ to denote the forward-degree of the (unique) node added at time $T_a(i)$ in $B_w(a, T_a(i))$. Recall that the forward-degree is the degree minus one. Define $\widehat{S}_a(i)$ as follows.

$$\widehat{S}_a(i) := d_a + \widehat{d}_a(1) + ... + \widehat{d}_a(i) - i, \qquad \hat{S}_a(0) = d_a. \tag{2.1}$$

For a connected graph H, the tree excess of H is denoted by $\mathrm{tx}(H)$, which is the maximum number of edges that can be deleted from H while still keeping it connected. By an abuse of notation, for a subset $W \subseteq V$, we denote by $\mathrm{tx}(W)$ the tree excess of the induced subgraph $G[W]$ of G on W. (If $G[W]$ is not connected, then $\mathrm{tx}(W) := \infty$.) Consider the growing balls $B_w(a, T_a(i))$ for $0 \le i \le I_a$ centered at a and let $X_a(i)$ be the tree excess of $B_w(a, T_a(i))$:

$$X_a(i) := \mathrm{tx}\,(\,B_w(a, T_a(i))\,).$$

We extend the definition of X_a to all the integer values by setting $X_a(i) = X_a(I_a)$ for all $i > I_a$.

The number of edges crossing the boundary of the ball $B_w(a, T_a(i))$ is denoted by $S_a(i)$. A simple calculation shows that

$$S_a(i) = \widehat{S}_a(i) - 2X_a(i). \tag{2.2}$$

We now consider a random graph $G(n, (d_i)_1^n)$ with i.i.d. rate one exponential weights on its edges, such that the degree sequence $(d_i)_1^n$ satisfies Condition 1.1. We let $m^{(n)}$ be the total degree defined by $m^{(n)} = \sum_{i=1}^n d_i = \sum_{k \ge 0} k u_k^{(n)}$.

One particularly useful property of the configuration model is that it allows one to construct the graph gradually, exposing the edges of the perfect matching one at a time. This way, each additional edge is uniformly distributed among all possible edges on the remaining (unmatched) half-edges. We have the following useful lemma.

Lemma 2.1. *For any $k \le \frac{m^{(n)}-n}{2}$, we have*

$$\mathbb{P}\left(2X_a(k) \ge x \mid \widehat{S}_a(k), I_a \ge k\right) \le \mathbb{P}\left(\mathrm{Bin}\left(\widehat{S}_a(k), \sqrt{\widehat{S}_a(k)/n}\right) \ge x \mid \widehat{S}_a(k)\right).$$

Proof. To prove this, we need the following intermediate result proved in [16, Lemma 3.2].

Lemma 2.2. *Let A be a set of m points, i.e., $|A| = m$, and let F be a uniform random matching of elements of A. For $e \in A$, we denote by $F(e)$ the point matched to e, and similarly for $X \subset A$, we write $F(X)$ for the set of points matched to X. Now let $X \subset A$, $k = |X|$, and assume $k \le m/2$. We have*

$$|X \cap F(X)| \le_{st} \mathrm{Bin}(k, \sqrt{k/m}).$$

Conditioning on all the possible degree sequences $\widehat{d}_a(1), \widehat{d}_a(2), \dots, \widehat{d}_a(k)$, with the property that $d_a + \sum_{1 \le i \le k} \widehat{d}_a(i) = \widehat{S}_a(k)$, the configuration model becomes equivalent to the following process: start from a and at each step $1 \le i \le k$, choose a vertex a_i of degree $\widehat{d}_a(i)+1$ uniformly at random from all the possible vertices of this degree outside the set $\{a, a_1, \dots, a_{i-1}\}$, choose a half-edge adjacent to a_i uniformly at random and match it with a uniformly chosen half-edge from the yet-unmatched half-edges adjacent to one of the nodes $a, a_1, \dots, a_{i-1}$. And at the end, after a_k has been chosen, take a uniform matching for all the remaining $(m^{(n)} - 2k)$ half-edges. Now the proof follows from Lemma 2.2 by the simple observation that, since $m^{(n)} - 2k \ge n$,

$$\mathbb{P}\left(\mathrm{Bin}\left(\widehat{S}_a(k), \sqrt{\widehat{S}_a(k)/m^{(n)} - 2k}\right) \ge x \mid \widehat{S}_a(k)\right) \\ \le \mathbb{P}\left(\mathrm{Bin}\left(\widehat{S}_a(k), \sqrt{\widehat{S}_a(k)/n}\right) \ge x \mid \widehat{S}_a(k)\right).$$

□

In the sequel, we will also need to consider the number of vertices of forward-degree at least two in the (growing) balls centered at a vertex $a \in V$. Thus, for $i \le I_a$, define

$$\gamma_a(i) := \sum_{\ell=1}^{i} \mathbb{1}[\widehat{d}_a(\ell) \ge 2] = |\Big\{b \in B_w(a, T_a(i)) : \; b \ne a \text{ and } d_b \ge 3\Big\}|, \tag{2.3}$$

and extend the definition to all integers by setting $\gamma_a(i) = \gamma_a(I_a)$ for all $i > I_a$. Note that $\gamma_a(0) = 0$ and $\gamma_a(i) = i$ if $d_{\min} \ge 3$.

Now define $\overline{T}_a(k)$ to be the first time where the ball centered at a has at least k nodes of forward-degree at least two. More precisely,

$$\overline{T}_a(i) \;\; := \;\; \min\Big\{T_a(\ell), \;\text{ for } \ell \text{ such that } \gamma_a(\ell) \ge k\Big\}. \tag{2.4}$$

The main idea of the proof of Theorem 1.2 consists in growing the balls around each vertex of the graph simultaneously so that the diameter becomes equal to twice the time when the last two balls intersect. In what follows, instead of taking a graph at random and then analyzing the balls, we use a standard coupling argument in random graph theory which allows to build the balls and the graph at the same time. We present this coupling in the next coming section.

2.1. *The exploration process.* Fix a vertex a in $G^*(n, (d_i)_1^n)$, and consider the following continuous-time exploration process. At time $t = 0$, we have a neighborhood consisting only of a, and for $t > 0$, the neighborhood is precisely $B_w(a,t)$. We now give an equivalent description of this process. This provides a more convenient way for analyzing the random variables which are crucial in our argument, e.g., $S_a(k)$. The idea is that instead of taking a graph at random and then analyzing the balls, the graph and the balls are built at the same time. We will consider a growing set of vertices denoted by B and a list L of yet unmatched half-edges in B. Recall that in the usual way of constructing a random graph with given degree sequence, we match half-edges amongst themselves uniformly at random. In the following, by a matching, we mean a pair of matched half-edges.

- Start with $B = \{a\}$, where a has d_a half-edges. For each half edge, decide (at random depending on the previous choices) if the half-edge is matched to a half-edge adjacent to a or not. Reveal the matchings consisting of those half-edges adjacent to a which are connected amongst themselves (creating self-loops at a) and assign weights independently at random to these edges. The remaining unmatched half-edges adjacent to a are stored in a list L. (See the next step including a more precise description of this first step.)
- Repeat the following exploration step as long as the list L is not empty.
- Given there are $\ell \geq 1$ half-edges in the current list, say $L = (h_1, \dots, h_\ell)$, let $\Psi \sim \mathrm{Exp}(\ell)$ be an exponential variable with mean ℓ^{-1}. After time Ψ select a half-edge from L uniformly at random, say h_i. Remove h_i from L and match it to a uniformly chosen half-edge in the entire graph excluding L, say h. Add the new vertex (connected to h) to B and reveal the matchings (and weights) of any of its half-edges whose matched half-edge is also in B. More precisely, let d be the degree of this new vertex and $2x$ the number of already matched half-edges in B (including the matched half-edges h_i and h). There is a total of $m - 2x$ unmatched half-edges, m being the total number of half-edges of the random graph G. Consider one of the $d - 1$ half-edges of the new vertex (excluding h which is connected to h_i); with probability $(\ell - 1)/(m - 2x - 1)$ it is matched with a half-edge in L and with the complementary probability it is matched with an unmatched half-edge outside L. In the first case, match it to a uniformly chosen half-edge of L and remove the corresponding half-edge from L. In the second case, add it to L. We proceed in the similar manner for all the $d - 1$ half-edges of the new vertex.

Let $B(a,t)$ and $L(a,t)$ be respectively the set of vertices and the list generated by the above procedure at time t, where a is the initial vertex.

Considering the usual configuration model and using the memoryless property of the exponential distribution, we have $B_w(a,t) = B(a,t)$ for all t. To see this, we can continuously grow the weights of the half-edges $h_1, \dots, h_\ell$ in L until one of their rate 1 exponential clocks fire. Since the minimum of ℓ i.i.d exponential variables with rate 1 is exponential with rate ℓ, this is the same as choosing uniformly a half-edge h_i after time Ψ (recall that by our conditioning, these ℓ half-edges do not pair within themselves). Note that the final weight of an edge is accumulated between the time of arrival of its first half-edge and the time of its pairing (except edges going back into B whose weights are revealed immediately). Then the equivalence follows from the memoryless property of the exponential distribution.

Note that $T_a(i)$ is the time of the i-th exploration step in the above continuous-time exploration process. Assuming $L(a, T_a(i))$ is not empty, at time $T_a(i+1)$, we match a uniformly chosen half-edge from the set $L(a, T_a(i))$ to a uniformly chosen half-edge among all other half-edges, excluding those in $L(a, T_a(i))$. Let $\mathcal{F}_t$ be the σ-field generated by the above process until time t. Given $\mathcal{F}_{T_a(i)}$, $T_a(i+1) - T_a(i)$ is an exponential random variable with rate $S_a(i)$ given by Equation (2.2) wich is equal to $|L(a, T_a(i))|$ the size of the list consisting of unmatched half-edges in $B(a, T_a(i))$. In other words,

$$\left(T_a(i+1) - T_a(i) \,|\, \mathcal{F}_{T_a(i)}\right) \overset{d}{=} \mathrm{Exp}(S_a(i)),$$

this is true since the minimum of k i.i.d. rate one exponential random variables is an exponential of rate k.

Recall that $I_a = \min\{i,\ S_a(i) = 0\} \leq n-1$, and set $S_a(i) = 0$ for all $I_a \leq i \leq n-1$. We now extend the definition of the sequence $\widehat{d}(i)$ to all the values of $i \leq n-1$, constructing a sequence $(\widehat{d}_a(i))_{i=1}^{n-1}$ which will coincide in the range $i \leq I_a$ with the sequence $\widehat{d}_a(i)$ defined in the previous subsection. We first note that in the terminology of the exploration process, the sequence $(\widehat{d}_a(i))_{i \leq I_a}$ can be constructed as follows. At time $T_a(i+1)$, the half-edge adjacent to the $i+1$-th vertex is chosen uniformly at random from the set of all the half-edges adjacent to a vertex out-side B, and $\widehat{d}(i+1)$ is the forward-degree of the vertex adjacent to this half-edge. Thus, the sequence $(\widehat{d}(i))_{i \leq I_a}$ has the following description.

Initially, associate to all vertex j a set of d_j half-edges (corresponding the set of half-edges outside B and L). At step 0, remove the half-edges corresponding to vertex a. Subsequently, at step $k \leq I_a$, choose a half-edge uniformly at random among all the remaining half-edges; if the half-edge is drawn from the node j's half-edges, then set $\widehat{d}_a(k) = d_j - 1$, and remove the node j and all of its half-edges. Obviously, this description allows to extend the definition of $\widehat{d}_a(i)$ to all the values of $I_a < i \leq n-1$. Indeed, if $I_a < n-1$, there are still half-edges at step $I_a + 1$, and we can complete the sequence $\widehat{d}_a(i)$ for $i \in [I_a + 1, n-1]$ by continuing the sampling described above. In this way, we obtain a sequence $(\widehat{d}_a(i))_{i=1}^{n-1}$ which coincides with the sequence defined in the previous section for $i \le I_a$.

We also extend the sequence $\widehat{S}_a(i)$ for $i > I_a$ thanks to (2.1). Recall that, we set $X_a(i) = X_a(I_a)$ for all $i > I_a$. It is simple to see that with these conventions, the relation (2.2) is not anymore valid for $i > I_a$ but we still have $S_a(i) \le \widehat{S}_a(i) - 2X_a(i)$ for all i.

The process $i \mapsto X_a(i)$ is non-decreasing in $i \in [1, n-1]$. Moreover, given $\mathcal{F}_{T_a(i)}$, the increment $X_a(i+1) - X_a(i)$ is stochastically dominated by the following binomial random variable

$$(2.5)\ X_a(i+1) - X_a(i) \le_{st} \mathrm{Bin}\left(\widehat{d}_a(i+1), \frac{(S_a(i)-1)^+}{m^{(n)} - 2(X_a(i)+i)}\right),$$

where $m^{(n)} = \sum_{i=1}^n d_i$. We recall here that for two real-valued random variables A and B, we say A is stochastically dominated by B and write $A \le_{st} B$ if for all x, we have $\mathbb{P}(A \ge x) \le \mathbb{P}(B \ge x)$. If C is another random variable, we write $A \le_{st} (B \,|\, C)$ if for all x, $\mathbb{P}(A \ge x) \le \mathbb{P}(B \ge x \,|\, C)$.

Note that if $i > I_a$, then $S_a(i) = 0$ and $X_a(i+1) - X_a(i) = 0$, so that (2.5) is still valid.

For $i < \frac{n}{2}$, we have

$$\begin{aligned} \frac{(S_a(i)-1)^+}{m^{(n)} - 2(X_a(i)+i)} &\le \frac{\widehat{S}_a(i) - 2X_a(i)}{m^{(n)} - 2(X_a(i)+i)} \\ &\le \frac{\widehat{S}_a(i)}{m^{(n)} - 2i} \le \frac{\max_{\ell \le i} \widehat{S}_a(\ell)}{n - 2i}. \end{aligned}$$

We conclude:

Lemma 2.3. *For $i < \frac{n}{2}$, we have*

$$(2.6)\qquad X_a(i) \le_{st} \mathrm{Bin}\left(\max_{\ell \le i} \widehat{S}_a(\ell) + i, \frac{\max_{\ell \le i} \widehat{S}_a(\ell)}{n-2i}\right).$$

An important ingredient in the proof will be the coupling of the forward-degrees sequence $\{\widehat{d}(i)\}$ to an i.i.d. sequence in the range $i \le \beta_n$, that we provide in the next subsection.

Recall that we defined α_n and β_n as follows (c.f. Equation (1.9))

$$\alpha_n = \lfloor \log^3 n \rfloor, \text{ and } \beta_n = \lfloor 3\sqrt{\frac{\mu}{\nu - 1} n \log n} \rfloor.$$

2.2. *Coupling the forward-degrees sequence $\widehat{d}_a(i)$.* We now present a coupling of the variables $\{\widehat{d}_a(1), ..., \widehat{d}_a(k)\}$ valid for $k \le \beta_n$, where β_n is defined in Equation (1.9), with an i.i.d. sequence of random variables, that we now define. Let $\Delta_n := \max_{i\in[1,n]} d_i$. Note that by Condition 1.1 (ii), we have $\Delta_n = O(n^{1/2-\epsilon})$.

Denote the order statistics of the sequence of degrees $(d_i^{(n)})$ by

$$d_{(1)}^{(n)} \le d_{(2)}^{(n)} \le \cdots \le d_{(n)}^{(n)} \,. \tag{2.7}$$

Define $\underline{m}^{(n)} := \sum_{i=1}^{n-\beta_n} d_{(i)}^{(n)}$ and let $\underline{\pi}^{(n)}$ be the size-biased empirical distribution with the β_n highest degrees in (2.7) removed, i.e.,

$$\underline{\pi}_k^{(n)} := \frac{\sum_{i=1}^{n-\beta_n} (k+1) 1\!\!1\left[d_{(i)}^{(n)} = k+1\right]}{\underline{m}^{(n)}}.$$

Similarly, define $\overline{m}^{(n)} := \sum_{i=(\beta_n+1)\Delta_n}^{n} d_{(i)}^{(n)}$ and let $\overline{\pi}^{(n)}$ be the size-biased empirical distribution with the $(\beta_n+1)\Delta_n$ lowest degrees in (2.7) removed, i.e.,

$$\overline{\pi}_k^{(n)} := \frac{\sum_{i=(\beta_n+1)\Delta_n}^{n} (k+1) 1\!\!1\left[d_{(i)}^{(n)} = k+1\right]}{\overline{m}^{(n)}}.$$

Note that by Condition 1.1, we have $\beta_n \Delta_n = o(n)$ which implies that both the distributions $\underline{\pi}^{(n)}$ and $\overline{\pi}^{(n)}$ converge to the size-biased distribution q defined in Equation (1.1) as n tends to infinity.

The following basic lemma, proved in [2, Lemma 4.1], shows that the forward-degree of the i-th vertex given the forward-degrees of all the previous vertices is stochastically between two random variables with lower and upper distributions $\overline{\pi}^{(n)}$ and $\underline{\pi}^{(n)}$ defined above, provided that $i \le \beta_n$. More precisely,

LEMMA 2.4. *For a uniformly chosen vertex a, we have for all $i \le \beta_n$,*

$$\underline{D}_i^{(n)} \le_{st} \left(\widehat{d}_a(i) \,|\, \widehat{d}_a(1), \dots, \widehat{d}_a(i-1)\right) \le_{st} \overline{D}_i^{(n)}\,, \tag{2.8}$$

where $\underline{D}_i^{(n)}$ (resp. $\overline{D}_i^{(n)}$) are i.i.d. with distribution $\underline{\pi}^{(n)}$ (resp. $\overline{\pi}^{(n)}$).
In particular, we have for all $i \le \beta_n$,

$$\sum_{k=1}^{i} \underline{D}_k^{(n)} \le_{st} \sum_{k=1}^{i} \widehat{d}_a(k) \le_{st} \sum_{k=1}^{i} \overline{D}_k^{(n)}.$$

3. Proof of the Upper Bound. In this section we present the proof of the upper bound for Theorem 1.2. Namely we prove that for any $\epsilon > 0$, with high probability for all vertices u and v which are in the same component (i.e., such that $\mathrm{dist}_w(u,v) < \infty$), we have

$$\mathrm{dist}_w(u,v) \quad \leq \quad \left(\frac{1}{\nu - 1} + \frac{2}{\Gamma(d_{\min})} \right) (1+\epsilon) \log n,$$

where $\Gamma(d_{\min})$ is defined in (1.4).

The proof will be based on the following two technical propositions. For the sake of readability, we postpone the proof of these two propositions to the end of this section.

The first one roughly says that for all u and v, the growing balls centered at u and v intersect w.h.p. provided that they contain each at least β_n nodes. More precisely,

Proposition 3.1. *We have w.h.p.*

$$\mathrm{dist}_w(u,v) \leq T_u(\beta_n) + T_v(\beta_n), \textit{ for all } u \textit{ and } v.$$

The above proposition shows that in proving the upper bound, it will be enough to control the random variable $T_u(\beta_n)$ for each node u in V. It turns out that in the range between α_n and β_n, in the cases $d_{\min} \geq 3$, $d_{\min} = 2$, and $d_{\min} = 1$, $T_u(k)$ have more or less the same behavior, namely, it takes time at most roughly half of the typical (weighted) distance to go from size α_n to β_n. More precisely,

Proposition 3.2. *For a uniformly chosen vertex u and any $\epsilon > 0$, we have*

$$\mathbb{P}\left(T_u(\beta_n) - T_u(\alpha_n) \geq \frac{(1+\epsilon)\log n}{2(\nu - 1)} \mid I_u \geq \alpha_n \right) = o(n^{-1}).$$

The conditioning $I_u \geq \alpha_n$ is here to ensure that the connected component which contains u has size at least α_n. In particular, note that one immediate corollary of the two above propositions is that two nodes whose connected components have size at least α_n are in the same component (necessarily the giant component), and that the two balls of size β_n centered at these two vertices intersect w.h.p.

Using the above two propositions, we are only left to understand $T_u(\alpha_n)$, and for this we will need to consider the cases $d_{\min} \geq 2$ and $d_{\min} = 1$ separately. Before going through the proof of the upper bound in these cases, we need one more result. Consider the exploration process started at a vertex a. We will need to find lower bounds for $S_a(k)$ in the range $1 \leq k \leq \alpha_n$. Recall that we defined $\gamma_a(k)$ as the number of nodes of forward-degree at least two in the growing balls centered at a, c.f. Equation (2.3) for the precise definition. These nodes are roughly all the ones which could contribute to the growth of the random variable $S_a(k)$. Now define the two following events.

$$\begin{aligned} R_a &:= \big\{S_a(k) \geq d_{\min} + \gamma_a(k), \text{ for all } 0 \leq k \leq \alpha_n - 1\big\}, \\ R'_a &:= \big\{S_a(k) \geq \gamma_a(k), \text{ for all } 0 \leq k \leq \alpha_n - 1\big\}. \end{aligned}$$

LEMMA 3.3. *Assume $d_a \geq 2$ and $\widehat{d}_a(i) \geq 1$ for all $1 \leq i \leq \alpha_n$. Then we have*

$$\mathbb{P}\left(R_a \mid \widehat{d}_a(1), \ldots \widehat{d}_a(n-1)\right) \geq 1 - o(\log^{10} n/n), \tag{3.1}$$

$$\mathbb{P}\left(R'_a \mid \widehat{d}_a(1), \ldots \widehat{d}_a(n-1)\right) \geq 1 - o(n^{-3/2}). \tag{3.2}$$

In particular, $\mathbb{P}(R_a) \geq 1 - o(\log^{10} n/n)$ and $\mathbb{P}(R'_a) \geq 1 - o(n^{-3/2})$.

PROOF. Since $\widehat{d}_a(i) \geq 1$, $\widehat{S}_a(k)$ is non-decreasing in k. We have for all $k \leq \alpha_n$,

$$d_{\min} + \gamma_a(k) \leq d_a + \gamma_a(k) \leq \widehat{S}_a(k) \leq \alpha_n \Delta_n = o(n), \tag{3.3}$$

and moreover, $\max_{k \leq \alpha_n} \widehat{S}_a(k) = \widehat{S}_a(\alpha_n)$. Since $d_a \geq 2$ and $S_a(k) = \hat{S}_a(k) - 2X_a(k)$, we have

$$\{X_a(\alpha_n) = 0\} \subset R_a, \quad \{X_a(\alpha_n) \leq 1\} \subset R'_a.$$

Note that the inequalities in (3.3) are true for any sequence such that $1 \leq \widehat{d}_a(i) \leq \Delta_n$. In particular, in the rest of the proof we condition on a realization of the sequence $\mathbf{d} = (d_a, \widehat{d}_a(1), \ldots \widehat{d}_a(n-1))$.

We distinguish two cases depending on whether or not $\widehat{S}_a(\alpha_n)$ is smaller than $3\alpha_n$. Denote this event by $\mathcal{Q}$ (and its complementary by $\mathcal{Q}^c$), i.e.,

$$\mathcal{Q} := \left\{ \widehat{S}_a(\alpha_n) < 3\alpha_n \right\}.$$

- **Case 1)** $\widehat{S}_a(\alpha_n) < 3\alpha_n$. Conditioning on $\mathcal{Q}$, by Lemma 2.3 we have

  $$X_a(\alpha_n) \leq_{st} \text{Bin}\left(4\alpha_n, \frac{3\alpha_n}{n - 2\alpha_n}\right).$$

  Thus, we have

  $$\begin{aligned} \mathbb{P}\left(X_a(\alpha_n) \geq 1 \mid \mathcal{Q}, \mathbf{d} \right) \leq \ & \mathbb{P}\left(\text{Bin}\big(4\alpha_n, \tfrac{3\alpha_n}{n-2\alpha_n}\big) \geq 1 \right) \leq O(\alpha_n^2/n), \\ \mathbb{P}\left(X_a(\alpha_n) \geq 2 \mid \mathcal{Q}, \mathbf{d} \right) \leq \ & \mathbb{P}\left(\text{Bin}\big(4\alpha_n, \tfrac{3\alpha_n}{n-2\alpha_n}\big) \geq 2 \right) \leq O(\alpha_n^4/n^2). \end{aligned}$$

We infer that

$$\begin{aligned}\mathbb{P}\left((R_a)^c \mid \mathcal{Q}, \mathbf{d} \right) &\leq O(\alpha_n^2/n),\\ \mathbb{P}\left((R_a')^c \mid \mathcal{Q}, \mathbf{d} \right) &\leq O(\alpha_n^4/n^2).\end{aligned}$$

- **Case 2)** $\widehat{S}_a(\alpha_n) \geq 3\alpha_n$. Note that in this case, we still have

$$\max_{k\leq\alpha_n} \widehat{S}_a(k) = \widehat{S}_a(\alpha_n) \leq \alpha_n\Delta_n = o(n).$$

Moreover, there exists $k \leq \alpha_n$ such that for all $\ell \leq k$, $\widehat{S}_a(\ell) < 3\alpha_n$ and $\widehat{S}_a(k+1) \geq 3\alpha_n$. Note that since we have conditioned on the degree sequence $\mathbf{d}$, the value of k is deterministic (k is not a random variable). Conditioning on the event $\mathcal{Q}^c$, we obtain by Lemma 2.3

$$\begin{aligned}X_a(k) &\leq_{st} \mathrm{Bin}\left(4\alpha_n, \frac{3\alpha_n}{n-2\alpha_n}\right), \qquad \text{and}\\ X_a(\alpha_n) &\leq_{st} \mathrm{Bin}\left(\alpha_n(\Delta_n+1), \frac{\alpha_n\Delta_n}{n-2\alpha_n}\right).\end{aligned} \tag{3.4}$$

By Condition 1.1 (ii), there exists a $\epsilon > 0$ such that $\Delta_n := O(n^{1/2-\epsilon})$. Let $m = \lceil 2\epsilon^{-1}\rceil$. Combining the last (stochastic) inequality together with the Chernoff's inequality applied to the right-hand side binomial random variable, we obtain

$$\begin{aligned}\mathbb{P}\left(X_a(\alpha_n) \geq m \mid \mathcal{Q}^c, \mathbf{d} \right) &\leq \mathbb{P}\left(\mathrm{Bin}\left(\alpha_n(\Delta_n+1), \frac{\alpha_n\Delta_n}{n-2\alpha_n}\right) \geq m\right)\\ &= O\left((\Delta_n^2\alpha_n^2/n)^m\right) = o(n^{-3}).\end{aligned}$$

We notice that for all $\ell > k$, we have $S_a(\ell) \geq 2\alpha_n - 2X_a(\alpha_n)$. Also for n large enough, we have $2\alpha_n - 2m \geq d_{\min} + \gamma_a(\ell)$. Therefore,

$$\begin{aligned}\left\{ X_a(k) = 0,\ X_a(\alpha_n) \leq m,\ \mathcal{Q}^c \right\} &\subset R_a \cap \mathcal{Q}^c, \qquad \text{and}\\ \left\{ X_a(k) \leq 1,\ X_a(\alpha_n) \leq m,\ Q_1^c \right\} &\subset R_a' \cap \mathcal{Q}^c.\end{aligned}$$

This in turn implies that

$$\begin{aligned}\mathbb{P}\left((R_a)^c \mid \mathcal{Q}^c, \mathbf{d} \right) &\leq \mathbb{P}\left(X_a(k) \geq 1 \mid \mathcal{Q}^c \right) + \mathbb{P}\left(X_a(\alpha_n) \geq m \mid \mathcal{Q}^c \right)\\ &\leq O(\alpha_n^2/n), \text{and,}\\ \mathbb{P}\left((R_a')^c \mid \mathcal{Q}^c, \mathbf{d} \right) &\leq \mathbb{P}\left(X_a(k) \geq 2 \mid \mathcal{Q}^c \right) + \mathbb{P}\left(X_a(\alpha_n) \geq m \mid \mathcal{Q}^c \right)\\ &\leq O(\alpha_n^4/n^2).\end{aligned}$$

In the above inequalities, we used (stochastic) Inequality (3.4) and Case 1 to bound the terms $\mathbb{P}\left(X_a(k) \geq 1 \mid \mathcal{Q}^c \right)$ and $\mathbb{P}\left(X_a(k) \geq 2 \mid \mathcal{Q}^c \right)$.

The lemma follows by the definition of α_n. □

We are now in position to provide the proof of the upper bound in the different cases depending on whether $d_{\min} \geq 2$, or $d_{\min} = 1$.

In the following, we will use the following property of the exponential random variables, without sometimes mentioning. If Y is an exponential random variable of rate μ, then for any $\theta < \mu$, we have $\mathbb{E}\left[e^{\theta Y}\right] = \frac{\mu}{\mu-\theta}$.

3.1. *Proof of the upper bound in the case $d_{\min} \geq 2$.* Consider the exploration process defined in Section 2.1 starting from a. Recall the definitions (2.3) and (2.4): $\gamma_a(i)$ is the number of nodes with forward-degree (strictly) larger than one until the i-th exploration step, and $\overline{T}_a(k)$ is the first time that the k-th node with the forward-degree (strictly) larger than one appears in the exploration process started at node a. We also define the sets

$$L_a(k) := \left\{\ell,\ \overline{T}_a(k) \leq T_a(\ell) < \overline{T}_a(k+1)\right\},$$

for $k \geq 0$, and let $n_a(k)$ be the smallest ℓ in $L_a(k)$. Clearly, we have $n_a(k) \geq k$ and

$$\gamma_a^{-1}(k) = L_a(k) = [n_a(k), n_a(k+1)-1].$$

Note that in the case $d_{\min} \geq 3$, we have $q_1 = \overline{\pi}_1^{(n)} = \underline{\pi}_1^{(n)} = 0, \gamma_a(k) = k$, $\overline{T}_a(k) = T_a(k)$ and $L_a(k) = \{k\}$. This case is also treated in [2]. However, our arguments bellow are still valid in this case.

For $x, y \in \mathbb{R}$, we denote $x \wedge y = \min(x, y)$. We will need the following lemma.

Lemma 3.4. *For a uniformly chosen vertex a, any $x > 0$, and any $\ell = O(\log n)$, we have*

$$\mathbb{P}\left(\, T_a(\alpha_n \wedge I_a) \geq x \log n + \ell \,\right) \leq o(n^{-1}) + o(e^{-d_{\min}(1-q_1)\ell}).$$

Proof. Recall that given the sequence $S_a(k)$, for $k < I_a$, the random variables $T_a(k+1) - T_a(k)$ are i.i.d. exponential random variables with mean $S_a(k)^{-1}$. First write

$$\begin{aligned} T_a(\alpha_n) &= \sum_{0 \leq j < \alpha_n} T_a(j+1) - T_a(j) \\ &\leq \sum_{k \leq K_n} \overline{T}_a(k+1) - \overline{T}_a(k), \end{aligned}$$

where K_n is the largest integer such that $n_a(K_n) \leq \alpha_n$.

We now show that for any $x > 0$ and $\ell = O(\log n)$,

$$\mathbb{P}\Big(T_a(\alpha_n) \geq x \log n + \ell, R_a \Big) = o(e^{-d_{\min}(1-q_1)\ell}). \tag{3.5}$$

Remark that a sum of a geometric (with parameter π) number of independent exponential random variables with parameter μ is distributed as an exponential random variable with parameter $(1-\pi)\mu$. For any $k \leq K_n$, we have:

$$\overline{T}_a(k+1) - \overline{T}_a(k) = \sum_{j \in L_a(k)} T_a(j+1) - T_a(j)$$

Assume R_a holds, then we have $S_a(j) \geq d_{\min} + k$ for all $j \in [n_a(k), n_a(k+1)-1] = L_a(k)$. Thus,

$$T_a(j+1) - T_a(j) \leq_{st} Y_{k,i} \sim \mathrm{Exp}(d_{\min} + k),$$

where $i = j - n_a(k) + 1$, and all the $Y_{k,i}$'s are independent. (For $i = 1, \ldots, |L_a(k)|$, $Y_{k,i}$ are exponential random variables with rate $d_{\min} + k$.)

For any positive t and θ, we obtain (for $\mathbf{d}_a := (d_a, \widehat{d}_a(1), \ldots, \widehat{d}_a(n-1))$)

$$\begin{aligned}
\mathbb{P}\Big(T_a(\alpha_n) - \overline{T}_a(1) \geq t, R_a \Big) &\leq \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}(R_a) \prod_{1 \leq k \leq K_n} e^{\theta(\overline{T}_a(k+1) - \overline{T}_a(k))} \mid \mathbf{d}_a \right]\right] e^{-\theta t} \\
&= \mathbb{E}\left[\prod_{1 \leq k \leq K_n} e^{\theta \sum_{i=1}^{|L_a(k)|} Y_{k,i}} \mathbb{P}\left(R_a \mid \mathbf{d}_a\right) \right] e^{-\theta t} \\
&\leq \prod_{1 \leq k \leq \alpha_n} \left(1 + \frac{\theta}{(d_{\min} + k)(1 - \underline{\pi}_1^{(n)}) - \theta} \right) e^{-\theta t},
\end{aligned}$$

where in the last inequality, we used the fact that the probability for a new node to have forward-degree one is at most $\underline{\pi}_1^{(n)}$, and so the length $|L_a(k)|$ is dominated by a geometric random variable with parameter $\underline{\pi}_1^{(n)}$. Taking $\theta = d_{\min}(1 - \underline{\pi}_1^{(n)})$ in the above inequality, we get

$$\begin{aligned}
\mathbb{P}\Big(T_a(\alpha_n) - \overline{T}_a(1) \geq t, R_a \Big) \quad &\leq \quad \prod_{1 \leq k \leq \alpha_n} \left(1 + \frac{d_{\min}(1 - \underline{\pi}_1^{(n)})}{(1 - \underline{\pi}_1^{(n)})k} \right) e^{-d_{\min}(1 - \underline{\pi}_1^{(n)})t} \\
&= \quad \prod_{1 \leq k \leq \alpha_n} (k + d_{\min})/k \; e^{-d_{\min}(1 - \underline{\pi}_1^{(n)})t} \\
&< \quad \alpha_n^3 e^{-d_{\min}(1 - \underline{\pi}_1^{(n)})t}.
\end{aligned}$$

In the same way, we can easily deduce that

$$\left(\overline{T}_a(1) \mid R_a\right) \leq_{\mathrm{st}} \mathrm{Exp}(d_{\min}(1-\underline{\pi}_1^{(n)})).$$

Let $t = x\log n + \ell$, and note that $\ell \leq C\log n$ for some large constant $C > 0$ (by assumption $\ell = O(\log n)$). Take any $0 < \epsilon < x(1-q_1)(C+x)^{-1}$; since for n sufficiently large, we have $\underline{\pi}_1^{(n)} \leq q_1 + \epsilon$, we obtain

$$\mathbb{P}\left(T_a(\alpha_n) \geq x\log n + \ell, R_a\right) \leq \frac{\alpha_n^3}{n^{d_{\min}(x(1-q_1-\epsilon)-\epsilon C)}} e^{-d_{\min}(1-q_1)\ell},$$

and (3.5) follows. Note that $x(1-q_1-\epsilon) - \epsilon C > 0$ by the choice of ϵ.

Assume now that the event $R'_a \cap R^c_a$ holds. Two cases can happen: either $I_a < \alpha_n$ or $I_a \geq \alpha_n$.

(Remark that in the case $d_{\min} \geq 3$, by Lemma 3.3 we have $\mathbb{P}(I_a \geq \alpha_n) \geq 1 - o(n^{-3/2})$. Indeed, for $d_{\min} \geq 3$, we have $\gamma_a(k) = k$ so that $R'_a \subseteq \{I_a \geq \alpha_n\} = \{S_a(k) \geq 1, \text{ for all } 0 \leq k \leq \alpha_n - 1\}$.)

If $I_a < \alpha_n$, then by the definition of R'_a, $0 = S_a(I_a) \geq \gamma_a(I_a)$, i.e., $\gamma_a(I_a) = 0$. In other words, the component of a is a union of cycles (or loops) having node a as a common node, and with total number of edges less than α_n. Hence, in this case, we have

$$\begin{aligned}
&\mathbb{P}\Big(R'_a,\, R^c_a, I_a < \alpha_n,\, T_a(I_a) \,\geq\, x\log n + \ell\Big)\\
&\qquad \leq \mathbb{P}\left(R^c_a \mid \mathbf{d}_a\right)\left(\sum_{0\leq k\leq \alpha_n} (\underline{\pi}_1^{(n)})^k \int_{x\log n+\ell}^{\infty} t^k \frac{e^{-t}}{k!}dt\right)\\
&\qquad \leq \log^{10} n/n\ (1-\underline{\pi}_1^{(n)})^{-1} \exp\Big(-(1-\underline{\pi}_1^{(n)})(x\log n + \ell)\Big) = o(n^{-1}),
\end{aligned}$$

where the last inequality follows from Inequality (3.1) in Lemma 3.3.
In the second case, when $I_a \geq \alpha_n$, let

$$\mathcal{Q} = R'_a \cap R^c_a \cap \left\{ I_a \geq \alpha_n \right\}.$$

If $\mathcal{Q}$ holds, by the definition of R'_a, we have $S_a(j) \geq k$ for all $j \in L_a(k)$. Thus,

$$T_a(j+1) - T_a(j) \leq_{st} Y_{k,i} \sim \mathrm{Exp}(k),$$

where $i = j - n_a(k) + 1$, and all the $Y_{k,i}$'s are independent. (for $i = 1, \dots, |L_a(k)|$, $Y_{k,i}$ are exponential random variables with rate k.) Hence, by the same argument as above, we have

$$\begin{aligned}\mathbb{P}\left(T_a(\alpha_n) - \overline{T}_a(2) \geq t, \mathcal{Q}\right) &\leq \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}(\mathcal{Q}) \prod_{2 \leq k \leq K_n} e^{\theta(\overline{T}_a(k+1) - \overline{T}_a(k))} \mid \mathbf{d}_a\right]\right] e^{-\theta t} \\ &\leq \mathbb{E}\left[\prod_{2 \leq k \leq K_n} e^{\theta \sum_{i=1}^{|L_a(k)|} Y_{k,i}} \mathbb{P}\left(R_a^c \mid \mathbf{d}_a\right)\right] e^{-\theta t} \\ &\leq \prod_{2 \leq k \leq \alpha_n} \left(1 + \frac{\theta}{k(1 - \underline{\pi}_1^{(n)}) - \theta}\right) e^{-\theta t} o\left(\frac{\log^{10} n}{n}\right),\end{aligned}$$

where the last inequality follows from Inequality (3.1) in Lemma 3.3. Thus, taking $\theta = 1 - \underline{\pi}_1^{(n)}$ gives

$$\begin{aligned}\mathbb{P}\left(T_a(\alpha_n) - \overline{T}_a(2) \geq t, \mathcal{Q},\right) &\leq \prod_{2 \leq k \leq \alpha_n} \left(1 + \frac{1}{k-1}\right) e^{-(1-\underline{\pi}_1^{(n)})t} o\left(\frac{\log^{10} n}{n}\right) \\ &\leq \alpha_n e^{-(1-\underline{\pi}_1^{(n)})t} o\left(\frac{\log^{10} n}{n}\right) = e^{-(1-\underline{\pi}_1^{(n)})t} o\left(\frac{\log^{13} n}{n}\right).\end{aligned}$$

Since $d_a \geq d_{\min}$, we can easily deduce that

$$\left(\overline{T}_a(2) \mid \mathcal{Q}\right) \leq_{\text{st}} \text{Exp}(d_{\min}(1 - \underline{\pi}_1^{(n)})) + \text{Exp}(1 - \underline{\pi}_1^{(n)}),$$

with these two exponential being independent and independent of $\mathcal{Q}$. Hence, we have

$$\begin{aligned}\mathbb{P}\left(\overline{T}_a(2) \geq t \mid \mathcal{Q}\right) &\leq \int_t^\infty d_{\min}(1 - \underline{\pi}_1^{(n)}) \left(e^{-(1-\underline{\pi}_1^{(n)})x} - e^{-d_{\min}(1-\underline{\pi}_1^{(n)})x}\right) \\ &\leq d_{\min} e^{-(1-\underline{\pi}_1^{(n)})t}.\end{aligned}$$

Thus,

$$\mathbb{P}\left(T_a(\alpha_n) \geq t, \mathcal{Q}\right) \leq e^{-(1-\underline{\pi}_1^{(n)})t} o\left(\frac{\log^{13} n}{n}\right).$$

Similar to the case where R_a holds (by fixing a constant ϵ small enough and using that for n sufficiently large $\underline{\pi}_1^{(n)} \leq q_1 + \epsilon$ for n large enough), we get

$$\mathbb{P}\left(T_a(\alpha_n) \geq x \log n + \ell, \mathcal{Q}\right) \leq o\left(\frac{\log^{13} n}{n^{1+(1-q_1-\epsilon)C}}\right) = o(n^{-1}).$$

Putting all the above arguments together, and considering the three disjoint cases $(R'_a)^c$ holds, R_a holds, and $R'_a \cap R_a^c$ holds (in which case either $I_a < \alpha_n$ or $I_a \geq \alpha_n$), we conclude that

$$\mathbb{P}\Big(T_a(\alpha_n \wedge I_a) \geq x \log n + \ell\Big) \leq o(e^{-d_{\min}(1-q_1)\ell}) + o(n^{-1}) + 1 - \mathbb{P}(R'_a).$$

To conclude the proof it suffices to use Lemma 3.3. □

We can now finish the proof of the upper bound in the case $d_{\min} \geq 2$. By Proposition 3.2, and Lemma 3.4 applied to $\ell = \frac{\log n}{d_{\min}(1-q_1)}$, we obtain that for a uniformly chosen vertex a and any $\epsilon > 0$, we have

$$\mathbb{P}\left(\infty > T_a(\beta_n \wedge I_a) \geq \left(\frac{1}{2(\nu-1)} + \frac{1}{d_{\min}(1-q_1)}\right)(1+\epsilon)\log n\right) = o(n^{-1}). \tag{3.6}$$

Indeed, the above probability can be bounded above by

$$\mathbb{P}\left(T_a(\alpha_n \wedge I_a) \geq \frac{1+\epsilon}{d_{\min}(1-q_1)} \log n\right) + \mathbb{P}\left(T_a(\beta_n) - T_a(\alpha_n) \geq \frac{1+\epsilon}{2(\nu-1)} \log n \mid I_a \geq \alpha_n\right),$$

and this is $o(n^{-1})$ by the above cited results.

Applying Equation (3.6) (and Lemma 3.4) and a union bound over a, we obtain

$$\mathbb{P}\left(\forall a,\ T_a(\beta_n \wedge I_a) \leq \left(\frac{1}{2(\nu-1)} + \frac{1}{d_{\min}(1-q_1)}\right)(1+\epsilon)\log n\right) = 1 - o(1). \tag{3.7}$$

Hence by Proposition 3.1, we have w.h.p. (for $d_{\min} \geq 2$)

$$\frac{\mathrm{diam}_w(G(n,(d_i)_1^n))}{\log n} \leq (1+\epsilon)\Big(\frac{1}{\nu-1} + \frac{1}{1-q_1}\mathbb{1}[d_{\min}=2] + \frac{2}{d_{\min}}\mathbb{1}[d_{\min}\geq 3]\Big).$$

This proves the bound on the diameter. To obtain the upper bound for the flooding time, we use Equation (3.7), and proceed as above by applying Proposition 3.2, and Lemma 3.4 applied to $\ell = \epsilon \log n$, to obtain that for a uniformly chosen vertex b, we have

$$\mathbb{P}\left(T_b(\beta_n \wedge I_b) \leq (\frac{1+\epsilon}{2(\nu-1)} + \epsilon)\log n\right) = 1 - o(1). \tag{3.8}$$

Clearly, Equations (3.7) and (3.8) imply that w.h.p.

$$\frac{\mathrm{flood}_w(a, G(n,(d_i)_1^n))}{\log n} \leq (1+\epsilon)\Big(\frac{1}{\nu-1} + \frac{1}{2(1-q_1)}\mathbb{1}[d_{\min}=2] + \frac{1}{d_{\min}}\mathbb{1}[d_{\min}\geq 3]\Big).$$

Similarly, (3.8) and Proposition 3.2 give an upper bound for (1.5).

The proof of the upper bound in this case is now complete.

3.2. *Proof of the upper bound in the case $d_{\min} = 1$.* In this section, we will need some results on the 2-core of the graph. Basic definitions and needed results are given in Appendix A.

We denote by $\mathcal{C}_a$ the event that a is connected to the 2-core of $G_n \sim G(n, (d_i)_1^n)$. It is well-known (c.f. Section A) that the condition $\nu > 1$ ensures that the 2-core of G_n has size $\Omega(n)$, w.h.p. We consider the graph $\tilde{G}_n(a)$ obtained from G_n by removing all vertices of degree one except a until no such vertices exist. If the event $\mathcal{C}_a$ holds, $\tilde{G}_n(a)$ consists of the 2-core of G_n and the unique path (empty if a belongs to the 2-core) from a to the 2-core. While, if the event $\mathcal{C}_a^c$ holds, then the graph $\tilde{G}_n(a)$ is the union of the 2-core of G_n and the isolated vertex a.

In order to bound the weighted distance between two vertices a and b, in what follows, we will consider two cases depending on whether both the vertices a and b are connected to the 2-core (i.e., the events $\mathcal{C}_a$ and $\mathcal{C}_b$ both hold), or both the vertices a and b belong to the same tree component of the graph. In the former case, we will show how to adapt the analysis we made in the case $d_{\min} \geq 2$ to this case. And in the latter case, we directly bound the diameter of all the tree components of the graph.

First note that $\tilde{G}_n(a)$ can be constructed by means of the configuration model with a new degree sequence $\tilde{d}$, c.f. Section A with $\tilde{d}_i \geq 2$ for all $i \neq a$. Consider the exploration process on the graph $\tilde{G}_n(a)$ and denote by $\widetilde{T}_a(i)$ the first time the ball $\widetilde{B}_w(a,t)$ in $\tilde{G}_n(a)$ reaches size $i+1$. Also, $\tilde{I}_a$ is defined similarly to I_a for the graph $\tilde{G}_n(a)$. We need the following lemma.

LEMMA 3.5. *For a uniformly chosen vertex a, any $x > 0$ and any $\ell = O(\log n)$, we have*

$$\mathbb{P}\left(\widetilde{T}_a(\alpha_n \wedge \tilde{I}_a) \geq x \log n + \ell\right) \leq o(n^{-1}) + o(e^{-(1-\lambda_*)\ell}).$$

PROOF. First note that if $\mathcal{C}_a$ does not hold, i.e., if a is not connected to the 2-core, we will have $\tilde{I}_a = 0$ (since $\tilde{d}_a = 0$), and there is nothing to prove. Now the proof follows the same lines as in the proof of Lemma 3.4. Note that conditional on $\mathcal{C}_a$, we have $\tilde{d}_a \geq 1$, hence by Lemma 3.3, we have $\mathbb{P}(R_a \mid \mathcal{C}_a, \tilde{\mathbf{d}}) \geq 1 - o(\log^{10} n/n)$, and similarly for R'_a. The only difference we have to highlight here, compared to the proof of Lemma 3.4, is that conditional on $R_a \cap \mathcal{C}_a$, we have $\tilde{S}_a(j) \geq 1 + k$ for all $j \in \tilde{L}_a(k)$, where $\tilde{S}_a(j)$ and $\tilde{L}_a(k)$ are defined in the same way as $S_a(j)$ and $L_a(k)$ for the graph $\tilde{G}(a)$. Take now $\theta = 1 - \underline{\tilde{\pi}}_1^{(n)}$ in the Chernoff bound, used in the proof of Lemma 3.4, where $\underline{\tilde{\pi}}^{(n)}$ is defined as $\underline{\pi}^{(n)}$ for the degree sequence $(\tilde{d}_1^{(n)}, ..., \tilde{d}_{\tilde{n}}^{(n)})$. The rest of the proof of Lemma 3.4 can then be easily adapted to obtain the same result provided we replace $2(1-q_1)$ by $(1-\lambda_*)$, which is precisely the statement of the current lemma. (Note that $\lambda_* = \tilde{q}_1$, c.f. Section A.) $\square$

By Proposition 3.2 applied to the graph $\tilde{G}_n(a)$ (note that $\tilde{\nu} = \nu$, c.f. see Section A), and Lemma 3.5 applied to $\ell = \frac{\log n}{1-\lambda_*}$, we obtain that for a uniformly chosen vertex a and any $\epsilon > 0$, we have

(3.9)
$$\mathbb{P}\left(\infty > \widetilde{T}_a(\beta_n \wedge \tilde{I}_a) \geq \left(\frac{1}{2(\nu-1)} + \frac{1}{1-\lambda_*}\right)(1+\epsilon)\log n\right) = o(n^{-1}).$$

Indeed the above probability can be bounded above by

$$\mathbb{P}\left(\widetilde{T}_a(\alpha_n \wedge \tilde{I}_a) \geq \frac{1+\epsilon}{1-\lambda_*}\log n\right) + \mathbb{P}\left(\widetilde{T}_a(\beta_n) - \widetilde{T}_a(\alpha_n) \geq \frac{1+\epsilon}{2(\nu-1)}\log n \mid \tilde{I}_a \geq \alpha_n\right),$$

and this is $o(n^{-1})$ by the above cited results.

Applying Equation 3.9 (and Lemma 3.5) and a union bound over a, we obtain

(3.10)
$$\mathbb{P}\left(\forall a,\ \widetilde{T}_a(\beta_n \wedge \tilde{I}_a) \leq \left(\frac{1}{2(\nu-1)} + \frac{1}{1-\lambda_*}\right)(1+\epsilon)\log n\right) = 1 - o(1).$$

To obtain the upper bound for the flooding time and the typical distance, we use Equation (3.10), and proceed as above by using Lemma 3.5 applied to $\ell = \epsilon \log n$, to obtain that for a uniformly chosen vertex b, we have

$$\mathbb{P}\left(\widetilde{T}_b(\beta_n \wedge \tilde{I}_b) \leq \left(\frac{1+\epsilon}{2(\nu-1)} + \epsilon\right)\log n\right) = 1 - o(1). \tag{3.11}$$

Clearly, Equation (3.10) together with Proposition 3.1 (since $\widetilde{T}_a(k) \geq T_a(k)$ for all k), imply the desired upper bound on the giant component of G_n and also on every components containing a cycle, i.e, connected to 2-core.

At this point, we are only left to bound the (weighted) diameter of the tree components. In particular, the following lemma concludes the proof.

LEMMA 3.6. *For two uniformly chosen vertices a, b, and any $\epsilon > 0$, we have*

$$\mathbb{P}\left(\frac{1+\epsilon}{1-\lambda_*}\log n < \mathrm{dist}_w(a,b) < \infty,\ \mathcal{C}_a^c,\ \mathcal{C}_b^c\right) = o(n^{-2}).$$

PROOF. We consider the graph $\tilde{G}_n(a,b)$ obtained from G_n by removing vertices of degree less than two except a and b until no such vertices exist. As shown in Section A, the random graph $\tilde{G}_n(a,b)$ can be still obtained by a configuration model, and has the same asymptotic parameters as the random graph $\tilde{G}_n(a)$ in the proof of the previous lemma. We denote again by $\tilde{d}$, the degree sequence of the random graph $\tilde{G}_n(a,b)$. Also, $\widetilde{T}_a$ and $\tilde{I}_a$ are defined similarly for the graph $\tilde{G}_n(a,b)$.

Trivially, we can assume $\tilde{d}_a = 1$ and $\tilde{d}_b = 1$, otherwise, either they are not in the same component and so $\mathrm{dist}_w(a,b) = \infty$, or one of them is in the 2-core, i.e., one of the two events $\mathcal{C}_a$ or $\mathcal{C}_b$ holds. Consider now the exploration process started at a until time k^* which is the first time either a node with forward-degree (strictly) larger than one appears or the time that the unique half-edge adjacent to b is chosen by the process. Let v^* be the node chosen at k^*. Note that $\tilde{d}_{v^*} = 1$ if and only if the half-edge incident to b is chosen at k^*. We have

$$\begin{aligned}
\mathbb{P}&\left(\frac{1+\epsilon}{1-\lambda_*}\log n < \mathrm{dist}_w(a,b) < \infty,,\ \mathcal{C}_a^c,\ \mathcal{C}_b^c\right) \\
&= \mathbb{P}\left(\widetilde{T}_a(k^*) > \frac{1+\epsilon}{1-\lambda_*}\log n, v^* = b, \tilde{d}_a = \tilde{d}_b = 1\right) \\
&\le \mathbb{P}\left(\widetilde{T}_a(k^*) > \frac{1+\epsilon}{1-\lambda_*}\log n, v^* = b \mid \tilde{d}_a = \tilde{d}_b = 1\right) \\
&= \mathbb{P}\left(\widetilde{T}_a(k^*) > \frac{1+\epsilon}{1-\lambda_*}\log n \mid \tilde{d}_a = \tilde{d}_b = 1\right) \\
&\qquad \times \mathbb{P}\left(\tilde{d}_{v^*} = 1 \mid \tilde{d}_{v^*} \ne 2, \tilde{d}_a = \tilde{d}_b = 1\right) = o(n^{-2}).
\end{aligned}$$

To prove the last equality above, first note $\mathbb{P}(\tilde{d}_{v^*} = 1 \mid \tilde{d}_{v^*} \ne 2, \tilde{d}_a = \tilde{d}_b = 1) = O(\frac{1}{n})$, this holds since $\nu = \tilde{\nu} > 1$ and v^* will be chosen before $o(n)$ steps, i.e., $k^* = o(n)$ (we will indeed prove something much stronger, that $k^* = O(\log n)$, c.f. Lemma 4.1 in the next section). And second, $\mathbb{P}\left(\widetilde{T}_a(k^*) > \frac{1+\epsilon}{1-\lambda_*}\log n \mid \tilde{d}_a = \tilde{d}_b = 1\right) = o(1/n)$, this follows by the same argument as in the proof of Lemma 3.5 applied to $\tilde{G}_n(a,b)$, and by setting $\ell = \frac{(1+\epsilon)\log n}{1-\lambda_*}$. $\square$

The proof of the upper bound in this case is now complete by taking a union bound over all a and b. We end this section by presenting the proof of Proposition 3.1 and Proposition 3.2 in the next subsection.

3.3. *Proof of Proposition 3.1 and Proposition 3.2.* We start this section by giving some preliminary results that we will need in the proof of Proposition 3.1 and Proposition 3.2.

LEMMA 3.7. *Let $\underline{D}_i^{(n)}$ be i.i.d. with distribution $\underline{\pi}^{(n)}$. For any $\eta < \nu$, there is a constant $\gamma > 0$ such that for n large enough we have*

$$\mathbb{P}\left(\underline{D}_1^{(n)} + \cdots + \underline{D}_k^{(n)} \le k\eta\right) \le e^{-\gamma k}. \tag{3.12}$$

PROOF. Let D^* be a random variable with distribution $\mathbb{P}(D^* = k) = q_k$ given in Equation (1.1) so that $\mathbb{E}[D^*] = \nu$. Let $\phi(\theta) = \mathbb{E}[e^{-\theta D^*}]$. For any $\epsilon > 0$, there exists $\theta_0 > 0$ such that for any $\theta \in (0, \theta_0)$, we have $\log \phi(\theta) < (-\nu + \epsilon)\theta$. By Condition 1.1 and the fact that $\beta_n \Delta_n = o(n)$, i.e., $\sum_{i=n-\beta_n+1}^{n} d_{(i)}^{(n)} = o(n)$, we have for any $\theta > 0$, $\lim_{n\to\infty} \underline{\phi}^{(n)}(\theta) = \phi(\theta)$, where $\underline{\phi}^{(n)}(\theta) = \mathbb{E}[e^{-\theta \underline{D}_1^{(n)}}]$. Also, for $\theta > 0$,

$$\mathbb{P}\left(\underline{D}_1^{(n)} + \cdots + \underline{D}_k^{(n)} \le \eta k\right) \quad \le \quad \exp\left(k\left(\theta\eta + \log \underline{\phi}^{(n)}(\theta)\right)\right).$$

Fix $\theta < \theta_0$ and let n be sufficiently large so that $\log \underline{\phi}^{(n)}(\theta) \le \log \phi(\theta) + \epsilon$. This yields

$$\begin{aligned}\mathbb{P}\left(\underline{D}_1^{(n)} + \cdots + \underline{D}_k^{(n)} \le \eta k\right) &\le \exp\left(k\left(\theta\eta + \log \phi(\theta) + \epsilon\theta\right)\right)\\ &\le \exp\left(k\theta\left(\eta - \nu + 2\epsilon\right)\right),\end{aligned}$$

which concludes the proof. □

The following lemma is the main step in the proof of both the propositions.

LEMMA 3.8. *For any $\epsilon > 0$, define the event*

$$R''_a := \left\{ S_a(k) \ge \frac{\nu - 1}{1+\epsilon} k, \ \textit{for all } \alpha_n \le k \le \beta_n \right\}.$$

For a uniformly chosen vertex a, we have $\mathbb{P}\left(R''_a \mid I_a \ge \alpha_n \right) \ge 1 - o(n^{-5})$.

Before giving the proof of this lemma, we recall the following basic result and one immediate corollary, for the proof see for example [27, Theorem 1].

LEMMA 3.9. *Let $n_1, n_2 \in \mathbb{N}$ and $p_1, p_2 \in (0,1)$. We have* $\mathrm{Bin}(n_1, p_1) \leq_{st} \mathrm{Bin}(n_2, p_2)$ *if and only if the following conditions hold*

(*i*) $n_1 \leq n_2$,
(*ii*) $(1-p_1)^{n_1} \geq (1-p_2)^{n_2}$.

In particular, we have

COROLLARY 3.10. *If $x \leq y = o(n)$, we have (for n large enough)*

$$x - \mathrm{Bin}(x, \sqrt{x/n}) \leq_{st} y - \mathrm{Bin}(y, \sqrt{y/n}).$$

PROOF. By the above lemma, it is sufficient to show $(x/n)^{x/2} \geq (y/n)^{y/2}$, and this is true because s^s is decreasing near zero (for $s < e^{-1}$). □

Now we go back to the proof of Lemma 3.8.

Proof of Lemma 3.8 By Lemmas 2.4 and 3.7, for any $\epsilon > 0$, $k \geq \alpha_n$ and n large enough, we have

$$\mathbb{P}\left(\widehat{d}_a(1) + ... + \widehat{d}_a(k) \leq \frac{\nu}{1+\epsilon/2}k\right) \leq e^{-\gamma k} = o(n^{-6}).$$

We infer that with probability at least $1 - o(n^{-6})$, for any $k \leq \beta_n$,

$$\frac{\nu - 1}{1+\epsilon/2}k < d_a + \widehat{d}_a(1) + ... + \widehat{d}_a(k) - k < (k+1)\Delta_n = o(n).$$

By the union bound over k, we have with probability at least $1 - o(n^{-5})$ that for all $\alpha_n \leq k \leq \beta_n$,

$$\frac{\nu - 1}{1+\epsilon/2}k < \widehat{S}_a(k) < (k+1)\Delta_n = o(n). \tag{3.13}$$

Hence in the remaining of the proof we can assume that the above condition is satisfied.

By Lemma 2.1, Corollary 3.10 and Inequalities (3.13), conditioning on $\widehat{S}_a(k)$ and $\{I_a \geq k\}$, we have

$$\begin{aligned}\left(S_a(k) \mid \{I_a \geq k\} \right) &\geq_{st} \frac{\nu-1}{1+\epsilon/2}k - \mathrm{Bin}\left(\frac{\nu-1}{1+\epsilon/2}k, \sqrt{\left(\frac{\nu-1}{1+\epsilon/2}k\right)/n}\right)\\ &\geq_{st} \frac{\nu-1}{1+\epsilon/2}k - \mathrm{Bin}\left(\nu k, \sqrt{\nu k/n}\right).\end{aligned}$$

By Chernoff's inequality, since $k\sqrt{k/n} = o(k/\sqrt{\alpha_n})$, we have

$$\mathbb{P}\left(\mathrm{Bin}(\nu k, \sqrt{\nu k/n}) \geq k/\sqrt{\alpha_n}\right) \leq \exp\left(-\frac{1}{3}k/\sqrt{\alpha_n}\right) = o(n^{-6}).$$

Moreover, conditioned on $\{I_a \geq k\}$, we have with probability at least $1 - o(n^{-6})$,

$$S_a(k) \geq \frac{\nu - 1}{1 + \epsilon/2}k - \frac{k}{\sqrt{\alpha_n}} \geq \frac{\nu - 1}{1+\epsilon}k,$$

for n large enough. Defining

$$R''_a(k) := \left\{ S_a(k) \geq \frac{\nu - 1}{1+\epsilon}k \right\} \text{ for } \alpha_n \leq k \leq \beta_n,$$

so that $R''_a = \bigcap_{k=\alpha_n}^{\beta_n} R''_a(k)$, we have

$$\mathbb{P}\left(R''_a(k) \mid I_a \geq k \right) \geq 1 - o(n^{-6}). \tag{3.14}$$

Thus, by using the fact that $R''_a(k-1) \subset \left\{ I_a \geq k \right\}$, we get

$$\begin{aligned}
\mathbb{P}\left(R''_a \mid I_a \geq \alpha_n \right) &= 1 - \mathbb{P}\left(\bigcup_{k=\alpha_n}^{\beta_n} R''_a(k)^c \mid I_a \geq \alpha_n \right) \\
&= 1 - \mathbb{P}\left(R''_a(\alpha_n)^c \cup \bigcup_{k=\alpha_n+1}^{\beta_n} \left(R''_a(k)^c \cap R''_a(k-1) \right) \mid I_a \geq \alpha_n \right) \\
&\geq 1 - \mathbb{P}\left(R''_a(\alpha_n)^c \cup \bigcup_{k=\alpha_n+1}^{\beta_n} \left(R''_a(k)^c \cap \left\{ I_a \geq k \right\} \right) \mid I_a \geq \alpha_n \right) \\
&\geq 1 - \sum_{k=\alpha_n}^{\beta_n} \mathbb{P}\left(R''_a(k)^c \ \mid \ I_a \geq k \right) \geq 1 - o(n^{-5}),
\end{aligned}$$

which concludes the proof.

We are now in position to provide the proof of both the propositions.

Proof of Proposition 3.1. Fix two vertices u and v. We can assume that $T_u(\beta_n), T_v(\beta_n) < \infty$, i.e., $I_u, I_v \geq \beta_n$, otherwise the statement of the proposition holds trivially for u and v. Note that $\mathrm{dist}_w(u,v) \leq T_u(\beta_n) + T_v(\beta_n)$ is equivalent to

$$B_w(u, T_u(\beta_n)) \cap B_w(v, T_v(\beta_n)) \neq \emptyset.$$

Hence, to prove the proposition we need to bound the probability that $B_w(v, T_v(\beta_n))$ does not intersect $B_w(u, T_u(\beta_n))$.

First consider the exploration process for $B_w(u,t)$ until reaching $t = T_u(\beta_n)$. We know by Lemma 3.8 that with probability at least $1 - o(n^{-5})$,

$$S_u(\beta_n) \geq (\nu - 1 - o(1))\beta_n.$$

(In other words, there are at least $(\nu-1-o(1))\beta_n$ half-edges in $B_w(u, T_u(\beta_n))$.)

Next, begin exposing $B_w(v,t)$. Each matching adds a uniform half-edge to the neighborhood of v. Therefore, the probability that $B_w(v, T_v(\beta_n))$ does not intersect $B_w(u, T_u(\beta_n))$ is at most

$$\left(1 - \frac{(\nu - 1 - o(1))\beta_n}{m^{(n)}}\right)^{\beta_n} \leq \exp[-(9 - o(1)) \log n] < n^{-4}$$

for large n (recall that $\beta_n^2 = \frac{9\lambda n \log n}{\nu - 1}$). The union bound over u and v completes the proof. $\square$

Proof of Proposition 3.2. Conditioning on the event R''_a defined in Lemma 3.8, we have for any $\alpha_n \leq k \leq \beta_n$,

$$T_a(k+1) - T_a(k) \leq_{st} Y_k \sim \mathrm{Exp}(S_a(k)) \leq_{st} \mathrm{Exp}\left(\frac{\nu - 1}{1 + \epsilon} k\right),$$

and all the Y_k's are independent.

Letting $s = \sqrt{\alpha_n}$, for n large enough we obtain that

$$\begin{aligned}
\mathbb{E}\left[e^{s(T_a(\beta_n) - T_a(\alpha_n))} \mid R''_a\right] &\leq \prod_{k=\alpha_n}^{\beta_n - 1} \left(1 + \frac{s}{\frac{(\nu-1)k}{1+\epsilon} - s}\right) \leq \prod_{k=\alpha_n}^{\beta_n - 1} \left(1 + \frac{s(1+2\epsilon)}{(\nu - 1)k}\right) \\
&\leq \exp\left[\frac{s(1+2\epsilon)}{\nu - 1} \sum_{k=\alpha_n}^{\beta_n - 1} \frac{1}{k}\right] \leq \exp\left[\frac{s(1+3\epsilon)\log n}{2(\nu - 1)}\right].
\end{aligned}$$

By Markov's inequality,

$$\begin{aligned}
&\mathbb{P}\left(T_a(\beta_n) - T_a(\alpha_n) \geq \frac{(1+4\epsilon)\log n}{2(\nu - 1)} \mid I_a \geq \alpha_n\right) \\
&\qquad \leq 1 - \mathbb{P}(R''_a) + \mathbb{E}\left[e^{s(T_a(\beta_n) - T_a(\alpha_n))} \mid R''_a\right] \exp\left(-\frac{s(1+4\epsilon)\log n}{2(\nu - 1)}\right) \\
&\qquad \leq \exp\left(-\frac{s\epsilon \log n}{2(\nu - 1)}\right) + o(n^{-5}) = o(n^{-1}),
\end{aligned}$$

which concludes the proof. $\square$

4. Proof of the Lower Bound. In this section we present the proof of the lower bound for Theorem 1.2. To prove the lower bound, it suffices to show that for any $\epsilon > 0$, there exists w.h.p. two vertices u and v such that

$$\mathrm{dist}_w(u, v) > (1 - \epsilon) \left(\frac{1}{\nu - 1} + \frac{2}{\Gamma(d_{\min})} \right) \log n.$$

As in the proof of the upper bound, the proof will be different depending whether $d_{\min} = 1$ or ≥ 2. So we start this section by proving some preliminary results, including some new notations and definitions, that we will need in the proof for these cases, and then divide the end of the proof into two cases.

Fix a vertex a in $G_n \sim G(n, (d_i)_1^n)$, and consider the exploration process, defined in Section 2.1. Recall that $\overline{T}_a(1)$ is the first time when the ball centered at a contains a vertex of forward-degree at least two (i.e., degree at least 3), c.f. Equation-Definition (2.4). To simplify the notation, we denote by C_a the ball centered at a containing exactly one node (possibly in addition to a) of degree at least 3:

$$C_a := B_w(a, \overline{T}_a(1)). \tag{4.1}$$

Note that there is a vertex u (of degree $d_u \geq 3$) in C_a which is not in any ball $B_w(a, t)$ for $t < \overline{T}_a(1)$ and we have $\max_{v \in C_a} \mathrm{dist}_w(a, v) = \mathrm{dist}_w(a, u)$. We define the degree of C_a as

$$\deg(C_a) = d_a + d_u - 2. \tag{4.2}$$

Remark that at time $\overline{T}_a(1)$ of the exploration process defined in Section 2.1 starting from a, we have at most $\deg(C_a)$ free half-edges, i.e., the list L contains at most $\deg(C_a)$ half-edges. (We have the equality if the tree excess until time $\overline{T}_a(1)$ is zero.) The following lemma shows that the size of C_a is relatively small.

Lemma 4.1. *Consider a random graph $G(n, (d_i)_1^n)$ where the degrees d_i satisfy Condition 1.1. There exists a constant $M > 0$, independent of n, such that w.h.p. for all the nodes a of the graph, we have $|C_a| \leq M \log n$.*

Proof. We consider the exploration process, defined in Section 2.1, starting from a uniformly chosen vertex a, and use the coupling of the forward-degrees we described in subsection 2.2. Recall in particular that each forward-degree $\widehat{d}(i)$ conditioned on the previous forward-degrees is stochastically larger than a random variable with distribution $\underline{\pi}^{(n)}$. This shows that, at each step of the exploration process, the probability of choosing a node of degree at most two (forward-degree one or zero) will be at most $\underline{\pi}_0^{(n)} + \underline{\pi}_1^{(n)} < 1 - \epsilon$, for some $\epsilon > 0$ (note that the asymptotic mean of $\underline{\pi}^{(n)}$ is ν, and by assumption $\nu > 1$). We conclude that there exists a constant $M > 0$ such that for all large n, $\mathbb{P}(|C_a| > M \log n) = o(n^{-1})$. The union bound over a completes the proof. $\square$

For two subsets of vertices $U, W \subset V$, the (weighted) distance between U and W is defined as usual,

$$\mathrm{dist}_w(U, W) := \min\left\{ \ \mathrm{dist}_w(u, w) \mid u \in U, \ w \in W \ \right\}.$$

For two nodes a, b, define the event $\mathcal{H}_{a,b}$ as

$$\mathcal{H}_{a,b} := \left\{ \frac{1-\epsilon}{\nu - 1} \log n < \mathrm{dist}_w(C_a, C_b) < \infty \right\}. \tag{4.3}$$

Note that $\frac{\log n}{\nu - 1}$ is the typical distance, so the left inequality in the definition of the above event means that C_a and C_b have the right typical distance in the graph (modulo a factor $(1-\epsilon)$). The right inequality simply means that a and b belong to the same connected component. The following proposition is the crucial step in the proof of the lower bound, the proof of which is postponed to the end of this section.

Proposition 4.2. *Consider a random graph $G(n, (d_i)_1^n)$ with i.i.d. rate one exponential weights on its edges. Suppose that the degree sequence $(d_i)_1^n$ satisfies Condition 1.1. Assume that the number of nodes with degree one satisfy $u_1^{(n)} = o(n)$, and let a and b be two distinct vertices such that $deg(C_a) = O(1)$, and $deg(C_b) = O(1)$. Then for all $\epsilon > 0$,*

$$\mathbb{P}\left(\mathcal{H}_{a,b}\right) = 1 - o(1).$$

Furthermore, the same result holds without the condition $deg(C_a) = O(1)$ (reps. $deg(C_b) = O(1)$) if the node a (reps. b) is chosen uniformly at random.

Note that in particular, Proposition 4.2 is still valid when a and b are chosen uniformly at random and hence provides a lower bound for (1.5).

Assuming the above proposition, we now show that

(i) If the minimum degree $d_{\min} \geq 2$, then there are pairs of nodes a and b of degree $d_{\min}$ such that $\mathcal{H}_{a,b}$ holds and in addition, the closest nodes to each with forward-degree at least two is at distance at least $(1-\epsilon)\log n/(d_{\min}(1-q_1))$ w.h.p., for all $\epsilon > 0$.

(*ii*) If the minimum degree $d_{\min} = 1$, then there are pairs of nodes of degree one such that $\mathcal{H}_{a,b}$ holds and in addition, the closest node to each which belongs to the 2-core is at least $(1-\epsilon)\log n/(1-\lambda_*)$ away w.h.p., for all $\epsilon > 0$.

This will finish the proof of the claimed lower bound.

4.1. *Proof of the lower bound in the case* $d_{\min} \geq 2$. Let V^* be the set of all vertices of degree $d_{\min}$. We call a vertex u in V^* *good* if $\overline{T}_u(1)$ is at least $\frac{1-\epsilon}{d_{\min}(1-q_1)}\log n$, i.e.,

$$\overline{T}_u(1) \geq \frac{1-\epsilon}{d_{\min}(1-q_1)}\log n,$$

and if in addition, $\deg(C_u) \leq K$ for a constant K chosen as follows. Let $\widehat{D}$ be a random variable with the size-biased distribution, i.e., $\mathbb{P}(\widehat{D} = k) = q_k$. The constant K is chosen in order to have with positive probability $\widehat{D} \leq K - d_{\min} + 1$ conditionned on the event that $\widehat{D} \geq 2$, i.e.,

$$y = y_K := \mathbb{P}\left(\widehat{D} \leq d_{\min} - 1 + K \mid \widehat{D} \geq 2\right) > 0. \tag{4.4}$$

It is easy to verify that such K exists since $\nu > 1$.

It will be convenient to consider the two events in the definition of good vertices separately, namely, for a vertex $u \in V^*$, define

$$\mathcal{E}_u := \left\{\overline{T}_u(1) \geq \frac{1-\epsilon}{d_{\min}(1-q_1)}\log n\right\}, \text{ and} \tag{4.5}$$

$$\mathcal{E}'_u := \{\deg(C_u) \leq K\}. \tag{4.6}$$

We note that in the case $d_{\min} \geq 3$, the event $\mathcal{E}_u$ for $u \in V^*$ is equivalent to having a weight greater than $\frac{1-\epsilon}{d_{\min}}\log n$ on all the $d_{\min}$ edges connected to u, and clearly, the two above events $\mathcal{E}'_u$ and $\mathcal{E}_u$ are independent (conditionally on $u \in V^*$, i.e., $d_u = d_{\min}$).

For $u \in V^*$, let A_u be the event that u is good, $A_u := \mathcal{E}_u \cap \mathcal{E}'_u$, and let Y be the total number of good vertices, $Y := \sum_u \mathbb{1}_{A_u}$. In the following, we first obtain a bound for the expected value of Y, and then use the second moment inequality to show that w.h.p. $Y = \Omega(n^\epsilon)$.

Consider the exploration process defined in Section 2.1, starting from a node $u \in V^*$. At the beginning, each step of the exploration process is an exponential with parameter $d_{\min}$ (since there are $d_{\min}$ yet-unmatched half-edges adjacent to the explored vertices). In each step, the probability that the new half-edge of the list L does not match to the other half-edge of L

is at least $1 - 1/n$. This follows by observing that there are at least n yet-unmatched half-edges (by $\nu > 1$), and by using Lemma 4.1 (which says that before $M \log n$ steps the exploration process meets a vertex of forward-degree at least two). By the forward-degree coupling arguments of Section 2.2, the probability that a new matched node be of forward-degree one is at least $\overline{\pi}_1^{(n)}$. This shows that, with probability at least $(1-1/n)\overline{\pi}_1^{(n)}$ the exploration process adds a new node of forward-degree one. This shows that the first step in the exploration process a vertex of forward-degree at least two is added will be stochastically bounded below by a geometric random variable of parameter $(1 - 1/n)\overline{\pi}_1^{(n)}$. Each step takes rate $d_{\min}$ exponential time. Therefore,

$$\begin{aligned}\mathbb{P}(\mathcal{E}_u) &= \mathbb{P}\left(\overline{T}_u(1) \geq \frac{1-\epsilon}{d_{\min}(1-q_1)} \log n\right) \\ &\geq \mathbb{P}\left(\mathrm{Exp}\left(d_{\min}\left(1-(1-1/n)\overline{\pi}_1^{(n)}\right)\right) \geq \frac{1-\epsilon}{d_{\min}(1-q_1)} \log n\right).\end{aligned}$$

In the last inequality we used the fact that a sum of a geometric (with parameter π) number of independent exponential random variables of rate μ is distributed as an exponential random variable of rate $(1-\pi)\mu$. Note that this in particular shows that

$$\mathbb{P}(\mathcal{E}_u) \geq (1-o(1)) \exp(-(1-\epsilon)\log n) = (1-o(1))n^{-1+\epsilon}.$$

By using the coupling arguments of Section 2.2 (and by using Lemma 4.1), to bound the forward-degrees from above (and below) by i.i.d. random variables having distributions $\overline{\pi}^{(n)}$ (and $\underline{\pi}^{(n)}$) and then using the fact that the asymptotic distributions of both $\overline{\pi}^{(n)}$ and $\underline{\pi}^{(n)}$ coincides with the size biased distribution $\{q_k\}$, we have $\mathbb{P}(\mathcal{E}'_u \mid \mathcal{E}_u) = (1 \pm o(1))y$. We conclude that $\mathbb{P}(A_u) = \mathbb{P}(\mathcal{E}'_u \mid \mathcal{E}_u)\mathbb{P}(\mathcal{E}_u) \geq (1 \pm o(1))y\, n^{-1+\epsilon}$.

This shows that

$$\mathbb{E}[Y] = \sum_{u \in V^*} \mathbb{P}(A_u) \geq (1 \pm o(1))\, y\, p_{d_{\min}}\, n^{\epsilon}.$$

Note that above, we used Condition 1.1 which implies that $|V^*| = (1 \pm o(1))p_{d_{\min}} n$.

We now show that $\mathrm{Var}(Y) = o(\mathbb{E}[Y]^2)$. Applying the Chebyshev inequality, this will show that $Y \geq \frac{2}{3} y p_{d_{\min}} n^{\epsilon}$ with high probability.

For any pair of vertices $u, v \in V^*$ such that $C_u \cap C_v = \emptyset$, conditioning on A_u does not have much effect on the asymptotic of the degree distribution (by Lemma 4.1 the size of each component C_u is at most $M \log n$), and hence, we deduce by the coupling argument of Section 2.2 that for u and v such that $C_u \cap C_v = \emptyset$,

$$\mathbb{P}(A_v \cap A_u) = (1 \pm o(1))\mathbb{P}(A_u)\mathbb{P}(A_v).$$

We infer

$$\begin{aligned}
&\mathrm{Var}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \mathbb{E}[\sum_{u,v \in V^*} \mathbb{1}_{A_u}\mathbb{1}_{A_v}] - \mathbb{E}[Y]^2 \\
&= \mathbb{E}[\sum_{u,v \in V^*:\, C_u \cap C_v \neq \emptyset} \mathbb{1}_{A_u}\mathbb{1}_{A_v} + \sum_{u,v \in V^*:\, C_u \cap C_v = \emptyset} \mathbb{1}_{A_u}\mathbb{1}_{A_v}] - \mathbb{E}[Y]^2 \\
&= \mathbb{E}[\sum_{u \in V^*} \mathbb{1}_{A_u} \sum_{v \in V^*:\, C_u \cap C_v \neq \emptyset} \mathbb{1}_{A_v} + \sum_{u,v \in V^*:\, C_u \cap C_v = \emptyset} \mathbb{1}_{A_u}\mathbb{1}_{A_v}] - \mathbb{E}[Y]^2 \\
&\leq (K+1)(M \log n)\ \mathbb{E}[Y] + \mathbb{E}[\sum_{u,v \in V^*:\, C_u \cap C_v = \emptyset} \mathbb{1}_{A_u}\mathbb{1}_{A_v}] - \mathbb{E}[Y]^2 \\
&= o(\mathbb{E}[Y]^2).
\end{aligned}$$

In the inequality above, we used Lemma 4.1 to bound w.h.p. the size of all C_w by $M \log n$ (for some large enough M) for any node w in the graph, and used the fact that if the event A_u holds, then there are at most K edges out-going from C_u. Each of the vertices v with the property that $C_u \cap C_v \neq \emptyset$ should be either already on C_u or connected with a path consisting only of vertices of degree two to C_u (in which case, this path should belong to C_v). A simple analysis then shows that the number of nodes v with the property that $C_u \cap C_v \neq \emptyset$ is bounded by $(K+1)M \log n$, and the inequality follows.

This finishes the proof of the fact that $Y \geq \frac{2}{3} p_{d_{\min}} y n^{\epsilon}$ with high probability.

We consider first the flooding time, and obtain the corresponding lower bound. Let Y' denote the number of good vertices that are at distance at most $\frac{1-\epsilon}{d_{\min}(1-q_1)} \log n + \frac{1-\epsilon}{\nu-1} \log n$ from a vertex a (chosen uniformly at random). It is clear that the lower bound follows by showing that $Y' < Y$ with high probability, i.e., $Y - Y' > 0$ w.h.p. To show this, we will bound the expected value of Y' and use Markov's inequality.

Since by Condition 1.1, V^* has size linear in n by applying Proposition 4.2, we obtain that for a uniformly chosen vertex $u \in V^*$, conditioning on A_u, we have $\mathbb{P}(\mathcal{H}_{a,u}) = 1 - o(1)$. Indeed, the two events $\mathcal{H}_{a,u}$ and $\mathcal{E}_u$ are independent, and conditioning on $\mathcal{E}'_u$ is the same as conditioning on $\deg(C_u) \leq K = O(1)$. Therefore, for a uniformly chosen vertex u in V^*, we have

$$\mathbb{P}\left(A_u \cap \mathcal{H}_{a,u}^c\right) = o(\mathbb{P}(A_u)),$$

where $\mathcal{H}_{a,u}^c$ denotes the complementary event of $\mathcal{H}_{a,u}$, i.e., the event that $\mathcal{H}_{a,u}$ does not occur. Thus, a straightforward calculation shows that $\mathbb{E}[Y'] = o(\mathbb{E}[Y]) = o(n^\epsilon)$. By Markov's inequality, we conclude that $Y' \leq \frac{1}{3} p_{d_{\min}} y n^\epsilon$ w.h.p., and hence $Y - Y'$ is w.h.p. positive. This implies the existence of a vertex u whose distance from a is at least $\left(\frac{1}{\nu-1} + \frac{1}{d_{\min}(1-q_1)}\right)(1-\epsilon)\log n$. Hence for any $\epsilon > 0$ we have w.h.p.

$$\begin{aligned}\mathrm{flood}_w(a, G_n) &\geq \max_{u \in V^*} \mathrm{dist}_w(a,u) \\ &\geq \left(\frac{1}{\nu-1} + \frac{1}{2(1-q_1)}\mathbb{1}[d_{\min}=2] + \frac{1}{d_{\min}}\mathbb{1}[d_{\min}\geq 3]\right)(1-\epsilon)\log n.\end{aligned}$$

We now turn to the proof of the lower bound for the (weighted) diameter of the graph. The proof will follows the same strategy as for the flooding time, but this time we need to consider the pairs of good vertices.

Let R denote the number of pairs of distinct good vertices. Recall we proved above that w.h.p. $Y \geq \frac{2}{3}\mathbb{E}[Y]$. Thus,

$$R = Y(Y-1) \geq \frac{2\mathbb{E}[Y]}{3}(\frac{2\mathbb{E}[Y]}{3} - 1) > \frac{1}{4}\mathbb{E}[Y]^2.$$

The probabilities that u and v are both good and $\mathcal{H}_{u,v}$ does not happen can be bounded as follows.

$$\begin{aligned}\mathbb{P}\left(A_u \cap A_v \cap \mathcal{H}_{u,v}^c\right) &= \mathbb{P}(A_u \cap A_v)\mathbb{P}(\mathcal{H}_{u,v}^c \mid A_u, A_v) \\ &= \mathbb{P}(A_u \cap A_v)\mathbb{P}\left(\mathcal{H}_{u,v}^c \mid \deg(C_u) \leq K, \deg(C_v) \leq K\right) \\ &\quad \text{(We used the independence of } \mathcal{H}_{u,v} \text{ and } \mathcal{E}_u \text{ and } \mathcal{E}_v\text{)} \\ &= o(\mathbb{P}(A_u \cap A_v)). \qquad (4.7)\end{aligned}$$

The last equality follows from Proposition 4.2, since C_u and C_v are of degree $O(1)$.

To conclude, consider R' the number of pairs of good vertices that are at distance at most $(1-\epsilon)(2\frac{\log n}{d_{\min}(1-q_1)} + \frac{\log n}{\nu - 1})$. By using Equation (4.7), we have $\mathbb{E}R' = o(\mathbb{E}[Y]^2)$. Applying Markov's inequality, we obtain that w.h.p. $R' \leq \frac{1}{6}(\mathbb{E}[Y])^2$, and thus, $R - R'$ is w.h.p positive. This implies that for any $\epsilon > 0$, we have w.h.p.

$$\begin{aligned}\mathrm{diam}_w(G_n) &\geq \max_{u,v \in V^*} \mathrm{dist}_w(u,v) \\ &\geq \left(\frac{1}{\nu-1} + \frac{1}{1-q_1}\mathbb{1}[d_{\min}=2] + \frac{2}{d_{\min}}\mathbb{1}[d_{\min}\geq 3]\right)(1-\epsilon)\log n.\end{aligned}$$

4.2. *Proof of the lower bound in the case* $d_{\min} = 1$. Consider the 2-core algorithm, and stop the process the first time the number of nodes of degree one drops below $n^{1-\epsilon/2}$. Let V^* be the set of all nodes of degree one at this time. We denote by $\tilde{G}_n(V^*)$ the graph constructed by configuration model on the set of remaining nodes (this is indeed the V^*-augmented 2-core). Observe that proving the lower bound on the graph $\tilde{G}_n(V^*)$ gives us the lower bound on G_n.

Since $|V^*| = o(n/\log n)$, and the 2-core has linear size in n, w.h.p. the degree sequence of $\tilde{G}_n(V^*)$ has the same asymptotic as the degree sequence in the 2-core of G_n (see Section A, Lemma A.2 for more details). In particular, we showed in Section A that for the size-biased degree sequence of the 2-core's degree distribution, we have $\tilde{q}_1 = \lambda_*$, and for its mean, we have $\tilde{\nu} = \nu$.

Repeating the coupling arguments of Section 2.2 and defining $\tilde{\overline{\pi}}^{(n)}$ (similar to the definition of $\overline{\pi}^{(n)}$) for the degree sequence of $\tilde{G}_n(V^*)$, we infer that $\tilde{\overline{\pi}}_1^{(n)} \to \lambda_*$.

Similar as before, call a vertex u in V^* *good* if both the events $\mathcal{E}_u$ and $\mathcal{E}'_u$ hold. Recall the definition of the two events

$$\mathcal{E}_u := \left\{ \tilde{\overline{T}}_u(1) \geq \frac{1-\epsilon}{1-\lambda_*} \log n \right\}, \text{ and} \tag{4.8}$$

$$\mathcal{E}'_u := \{\deg(C_u) \leq K\}. \tag{4.9}$$

Here the constant $K \geq 2$ is chosen with the property that $\tilde{q}_K > 0$ ($\tilde{q}$ is the size-biased probability mass function corresponding to the 2-core, c.f. Section A), and $\tilde{\overline{T}}_u$ is defined similarly to $\overline{T}_u$ for the graph $\tilde{G}_n(V^*)$.

Consider the exploration process starting from a node $u \in V^*$. At the beginning, each step of the exploration process is an exponential of rate one, and the probability that each new matched node be of forward-degree exactly one is at least $\tilde{\overline{\pi}}_1^{(n)}$. Similar to the case of $d_{\min} = 2$, we obtain

$$\begin{aligned}
\mathbb{P}(A_u) &\geq (1 \pm o(1))\tilde{q}_K \, \mathbb{P}\left(\mathrm{Exp}\left(1 - \tilde{\overline{\pi}}_1^{(n)}\right) \geq 1 - \epsilon(1-\lambda_*)\log n\right) \\
&= (1 \pm o(1))\tilde{q}_K \exp(-(1-\epsilon)\frac{1-\lambda_*}{1-\tilde{\overline{\pi}}_1^{(n)}}\log n) \\
&= (1 \pm o(1))\tilde{q}_K \, n^{-1+\epsilon}.
\end{aligned}$$

This shows that

$$\mathbb{E}[Y] = \sum_{u \in V^*} \mathbb{P}(A_u) \geq n^{1-\epsilon/2}(1 \pm o(1))\tilde{q}_K \, n^{-1+\epsilon} = (1 \pm o(1))\tilde{q}_K n^{\epsilon/2},$$

Similarly, we obtain that $\mathrm{Var}(Y) = o(\mathbb{E}[Y]^2)$, and the rest of the proof follows similarly to the precedent case by using Proposition 4.2 for $\tilde{G}_n(V^*)$. Note that in $\tilde{G}_n(V^*)$, the number of vertices of degree one is $o(n) = o(|\tilde{G}_n(V^*)|)$ and thus, Proposition 4.2 can be applied.

At the present we are only left to prove Proposition 4.2.

4.3. *Proof of Proposition 4.2.* In this section we present the proof of Proposition 4.2. It is shown in [23, 29] that the giant component of a random graph $G(n,(d_i)_1^n)$ for $(d_i)_1^n$ satisfying Condition 1.1 contains w.h.p. all but $o(n)$ vertices (since $\nu > 1$ and $u_0^{(n)} + u_1^{(n)} = o(n)$). This immediately shows that $\mathbb{P}(\mathrm{dist}_w(C_a, C_b) < \infty) = 1 - o(1)$. Define $t_n := \frac{1-\epsilon}{2(\nu-1)} \log n$. So to prove the proposition, we need to prove that $\mathrm{dist}_w(C_a, C_b)$ is lower bounded by t_n w.h.p. in the case where either $\deg(C_a) = O(1)$ (reps. $\deg(C_b = O(1))$ or a (reps. b) is chosen uniformly at random.

In the case where a is chosen uniformly at random, it is easy to deduce, by using Markov's inequality, that we have w.h.p. $\deg(C_a) \leq \log n$. Indeed, this is true since $\deg(C_a)$ is asymptotically distributed as $\left(D + \widehat{D} - 1 \mid \widehat{D} \geq 2\right)$, where $\widehat{D}$ is a random variable with the size-biased distribution, and D is independent of $\widehat{D}$ with the degree distribution $\{p_k\}$. (To show this, one can use the coupling argument of Section 2.2 to bound $\deg(C_a)$ stochastically from above.) And, since this latter random variable has finite moment (by Condition 1.1), by applying Markov's inequality, we obtain w.h.p. $\deg(C_a) \leq \log n$. This shows that in all cases stated in the proposition, we can assume that $\deg(C_a) \leq \log n$ and $\deg(C_b) \leq \log n$.

We now consider the exploration process defined in Section 2.1 starting from C_a, i.e., we start the exploration process with $B = C_a$, and apply the steps one and two of the process. In a similar way we defined $T_a(i)$, we define $T_{C_a}(i)$ to be the time of the i-th step in this continuous-time exploration process. Similarly, let $\widehat{d}_{C_a}(i)$ be the forward-degree of the vertex added at i-th exploration step for all $i \geq 1$, and define

$$\widehat{S}_{C_a}(i) := \deg(C_a) + \widehat{d}_{C_a}(1) + ... + \widehat{d}_{C_a}(i) - i, \tag{4.10}$$

and define $S_{C_a}(i)$ similarly, so that we have $S_{C_a}(i) \leq \widehat{S}_{C_a}(i)$. Note that $T_{C_a}(i)$ obviously satisfies

$$T_{C_a}(i+1) - T_{C_a}(i) = \mathrm{Exp}\left(S_{C_a}(i)\right) \geq_{st} Y_i \sim \mathrm{Exp}\left(\widehat{S}_{C_a}(i)\right),$$

where the random variables Y_i are all independent.

Also, we infer (by Lemma 2.4) that

$$\widehat{S}_{C_a}(i) \leq_{st} \log n + \sum_{j=1}^{i} \overline{D}_j^{(n)} - i, \tag{4.11}$$

where $\overline{D}_j^{(n)}$ are i.i.d with distribution $\overline{\pi}^{(n)}$.

Let $\overline{\nu}^{(n)}$ be the expected value of $\overline{D}_1^{(n)}$ which is

$$\overline{\nu}^{(n)} := \sum_k k \overline{\pi}_k^{(n)},$$

and define $z_n = \sqrt{n/\log n}$. We will show later that the two growing balls in the exploration processes started from C_a and C_b, for a and b as in the proposition, will not intersect w.h.p. provided that they are of size less than z_n. We now prove that $T_{C_a}(z_n) \geq t_n$ with high probability.
For this, let us define

$$T'(k) \sim \sum_{i=1}^{k} \mathrm{Exp}\left(\log n + \sum_{j=1}^{i} \overline{D}_j^{(n)} - i\right),$$

where all the exponential variables in the above sum are independent, such that by the above arguments, we have

$$T_{C_a}(z_n) \geq_{st} T'(z_n).$$

We need the following lemma. (We define $\mathrm{Exp}(s) := +\infty$ for $s \leq 0$.)

LEMMA 4.3. *Let $X_1, ..., X_t$ be a random process adapted to a filtration $\mathcal{F}_0 = \sigma[\emptyset], \mathcal{F}_1, ..., \mathcal{F}_t$, and let $\mu_i = \mathbb{E}X_i$, $\mathbf{\Sigma}_i = X_1 + ... + X_i$, $\Lambda_i = \mu_1 + ... + \mu_i$. Let $Y_i \sim \mathrm{Exp}(\Sigma_i)$, and $Z_i \sim \mathrm{Exp}(\Lambda_i)$, where all exponential variables are independent. Then we have*

$$Y_1 + ... + Y_t \geq_{st} Z_1 + ... + Z_t.$$

PROOF. By Jensen's inequality, it is easy to see that for positive random variable X, we have

$$\mathrm{Exp}(X) \geq_{st} \mathrm{Exp}(\mathbb{E}X).$$

Then by induction, it suffices to prove that for a pair of random variables X_1, X_2 we have $Y_1 + Y_2 \geq_{st} Z_1 + Z_2$. We have

$$\begin{aligned}
\mathbb{P}(Y_1 + Y_2 > s) &= \mathbb{E}_{X_1}[\mathbb{P}(Y_1 + Y_2 > s | X_1)] \\
&\geq \mathbb{E}_{X_1}[\mathbb{P}(\mathrm{Exp}(X_1) + \mathrm{Exp}(X_1 + \mu_2) > s)] \\
&\geq \mathbb{P}(Z_1 + Z_2 > s).
\end{aligned}$$

□

We infer by Lemma 4.3,

$$T'(z_n) \geq_{st} \sum_{i=0}^{z_n} \mathrm{Exp}\left(\log n + (\overline{\nu}^{(n)} - 1)i\right) =: T^*(z_n),$$

where all exponential variables are independent.

We now let $b_n := \log n - (\overline{\nu}^{(n)} - 1)$, so that we have

$$\begin{aligned}\mathbb{P}(T^*(z_n) \leq t) &\leq \int_{\sum x_i \leq t} e^{-\sum_{i=1}^{z_n}((\overline{\nu}^{(n)}-1)i+b_n)x_i} dx_1 ... dx_{z_n} \prod_{i=1}^{z_n}((\overline{\nu}^{(n)} - 1)i + b_n) \\ &= \int_{0 \leq y_1 \leq \cdots \leq y_{z_n} \leq t} e^{-(\overline{\nu}^{(n)}-1)\sum_{i=1}^{z_n} y_i} e^{-b_n y_{z_n}} dy_1 ... dy_{z_n} \prod_{i=1}^{z_n}((\overline{\nu}^{(n)} - 1)i + b_n),\end{aligned}$$

where $y_k = \sum_{i=0}^{k-1} x_{z_n - i}$. Letting y play the role of y_{z_n}, and accounting for all permutations over $y_1, ..., y_{z_n-1}$ (giving each such variable the range $[0, y]$), we obtain

$$\begin{aligned}\mathbb{P}(T^*(z_n) \leq t) &\leq (\overline{\nu}^{(n)} - 1)^{z_n} \frac{\prod_{i=1}^{z_n}(i + \frac{b_n}{\overline{\nu}^{(n)}-1})}{(z_n - 1)!} \\ &\qquad \cdot \int_0^t e^{-(\overline{\nu}^{(n)}-1+b_n)y} \left(\int_{[0,y]^{z_n-1}} e^{-(\overline{\nu}^{(n)}-1)\sum_{i=1}^{z_n-1} y_i} dy_1 ... dy_{z_n-1}\right) dy \\ &= z_n \frac{\prod_{i=1}^{z_n}(i + \frac{b_n}{\overline{\nu}^{(n)}-1})}{z_n!} (\overline{\nu}^{(n)} - 1) \\ &\qquad \cdot \int_0^t e^{-(\overline{\nu}^{(n)}-1+b_n)y} \left(\prod_{i=1}^{z_n-1} \int_0^y (\overline{\nu}^{(n)} - 1) e^{-(\overline{\nu}^{(n)}-1)y_i} dy_i\right) dy \\ &= z_n \prod_{i=1}^{z_n}(1 + \frac{b_n}{(\overline{\nu}^{(n)} - 1)i})(\overline{\nu}^{(n)} - 1) \\ &\qquad \cdot \int_0^t e^{-(\overline{\nu}^{(n)}-1+b_n)y} \left(1 - e^{-(\overline{\nu}^{(n)}-1)y}\right)^{z_n-1} dy \\ &\leq c z_n^{\frac{b_n}{\overline{\nu}^{(n)}-1}+1} (\overline{\nu}^{(n)} - 1) \int_0^t e^{-(\overline{\nu}^{(n)}-1+b_n)y} \left(1 - e^{-(\overline{\nu}^{(n)}-1)y}\right)^{z_n-1} dy,\end{aligned}$$

where $c > 0$ is an absolute constant. Recall that $t_n = \frac{1-\epsilon}{2(\nu-1)} \log n$, and $z_n = \sqrt{n/\log n}$. Now we use the fact that $\left(1 - e^{-(\overline{\nu}^{(n)}-1)y}\right)^{z_n-1} \leq e^{-n^{\alpha}}$, for some $\alpha > 0$ and for all $0 \leq y \leq t_n$. We infer

$$\mathbb{P}(T^*(z_n) \leq t_n) \leq c(\overline{\nu}^{(n)} - 1) z_n^{\frac{b_n}{\overline{\nu}^{(n)}-1}+1} \int_0^{t_n} e^{-n^{\alpha}} dy = o(n^{-4}),$$

since $b_n = O(\log n)$. Hence, we have w.h.p.

$$|B_w(C_a, t_n)| \le z_n.$$

(Here naturally, for $W \subseteq V$, we let $B_w(W, t) = \{b, \text{ such that } \mathrm{dist}_w(W, b) \le t\}$.)

Similarly for b, and exposing $B_w(C_b, t_n)$, again w.h.p we obtain a set of size at most z_n. Now remark that, because each matching is uniform among the remaining half-edges, the probability of hitting $B_w(C_a, t_n)$ is at most $\widehat{S}_{C_a}(z_n)/n$.

Let $\epsilon_n := \log\log n$. By Markov's inequality we have

$$\begin{aligned} \mathbb{P}\left(\widehat{S}_{C_a}(z_n) \ge z_n\epsilon_n\right) &\le \mathbb{E}\widehat{S}_{C_a}(z_n)/z_n\epsilon_n \\ &= \frac{K + (\overline{\nu}^{(n)} - 1)(z_n + \lambda_n)}{z_n\epsilon_n} = o(1). \end{aligned}$$

We conclude

$$\begin{aligned} &\mathbb{P}\Big(B_w(C_a, t_n) \cap B_w(C_b, t_n) \ne \emptyset \Big) \\ &\quad\le \mathbb{P}\Big(|B_w(C_a, t_n)| > z_n \Big) + \mathbb{P}\Big(|B_w(C_b, t_n)| > z_n \Big) \\ &\qquad + \mathbb{P}\left(\widehat{S}_{C_a}(z_n) \ge z_n\epsilon_n\right) + \epsilon_n z_n^2/n = o(1). \end{aligned}$$

This completes the proof of Proposition 4.2.

APPENDIX A: STRUCTURE OF THE 2-CORE

The k-core of a given graph G is the largest induced subgraph of G with minimum vertex degree at least k. The k-core of an arbitrary finite graph can be found by removing vertices of degree less than k, in an arbitrary order, until no such vertices exist.

Consider now a random graph $G_n \sim G^*(n, (d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.1. In the process of constructing a random graph G_n by matching the half-edges, the k-core can be found by successively removing the half-edge of a node of degree less than k followed by removing a uniformly random half-edge from the set of all the remaining half-edges until no such vertices (of degree less than k) remain. What remains at this time is the k-core. Since these half-edges are unexposed, the k-core edge set is uniformly random conditional on the k-core half-edge set. Let $k \ge 2$ be a fixed integer, and $\mathrm{Core}_k^{(n)}$ be the k-core of the graph $G_n \sim G^*(n, (d_i)_1^n)$. For integers $l \ge 0$ and $0 \le r \le l$, let π_{lr} denote the binomial probabilities

$$\pi_{lr}(p) = \mathbb{P}\left(\mathrm{Bin}(l, p) = r\right) = \binom{l}{r} p^r (1-p)^{l-r}.$$

We further define the functions

$$h(p) := \sum_{r=k}^{\infty} \sum_{l=r}^{\infty} r p_l \pi_{lr}(p), \text{ and } h_1(p) := \sum_{r=k}^{\infty} \sum_{l=r}^{\infty} p_l \pi_{lr}(p).$$

THEOREM A.1 (Janson, Luczak [21]). *Consider a random graph $G(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.1. Let $k \geq 2$ be fixed, and let $\mathrm{Core}_k^{(n)}$ be the k-core of $G(n,(d_i)_1^n)$. Let $\hat{p}$ be the largest $p \leq 1$ such that $\mu p^2 = h(p)$. Assume $\hat{p} > 0$, and further suppose that $\hat{p}$ is not a local maximum point of the function $h(p) - \mu p^2$. Then*

$$v\left(\mathrm{Core}_k^{(n)}\right)/n \xrightarrow{p} h_1(\hat{p}) > 0, \quad v_j\left(\mathrm{Core}_k^{(n)}\right)/n \xrightarrow{p} \sum_{l=j}^{\infty} p_l \pi_{lj}(\hat{p})$$

for $j \geq k$, *and* $e\left(\mathrm{Core}_k^{(n)}\right)/n \xrightarrow{p} \mu\hat{p}^2/2$.

From now on, we consider the case $k = 2$, and denote by $\tilde{G}$ the 2-core of a graph G. In particular applying Theorem A.1 to the case $k = 2$, we have $h(\hat{p}) := \mu\hat{p} - \sum_l l\ p_l\ \hat{p}(1-\hat{p})^{l-1} = \mu\hat{p}\ (1 - G_q(1-\hat{p}))$. Recall from Theorem A.1 that we have to solve the equation $\mu\hat{p}^2 = h(\hat{p})$, thus, we obtain $1 - \hat{p} = G_q(1-\hat{p})$, and so $\hat{p} = 1 - \lambda$.

By Theorem 10 in [28], the graph $\tilde{G}_n$ obtained from G_n has the same distribution as a random graph constructed by the configuration model on $\tilde{n}$ nodes with a degree sequence $\tilde{d}_1^{(n)}, ..., \tilde{d}_{\tilde{n}}^{(n)}$ satisfying the following properties,

$$\begin{aligned} \tilde{n}/n \xrightarrow{p} h_1(1-\lambda) &= 1 - G_p(\lambda) - (1-\lambda)G_p'(\lambda) \\ &= 1 - G_p(\lambda) - \mu\lambda(1-\lambda) > 0, \end{aligned}$$

and

$$\begin{aligned} |\{i, \tilde{d}_i^{(n)} = j\}|/n &\xrightarrow{p} \sum_{\ell=j}^{\infty} p_\ell \binom{\ell}{j} (1-\lambda)^j \lambda^{\ell-j}, \ j \geq 2, \\ \sum_i \tilde{d}_i^{(n)}/n &\xrightarrow{p} \mu(1-\lambda)^2. \end{aligned}$$

It follows that the sequence $\{\tilde{d}_1^{(n)}, ..., \tilde{d}_{\tilde{n}}^{(n)}\}$ satisfies also the Condition 1.1 for some probability distribution $\tilde{p}_k$ with mean $\tilde{\mu}$ (which can be easily calculated from the two above properties).

Let $\tilde{q}$ be the size-biased probability mass function corresponding to $\tilde{p}$. We now show that $\tilde{q}$ and q have the same mean. Indeed, denoting by $\tilde{\nu}$ the mean of $\tilde{q}$, we see that $\tilde{\nu}$ is given by

$$\begin{aligned}
\tilde{\nu} &:= \sum_k k\tilde{q}_k = \frac{1}{\tilde{\mu}}\sum_k k(k-1)\tilde{p}_k \\
&= \frac{\sum_{k\geq 2} k(k-1)\sum_{\ell\geq k} p_\ell \binom{\ell}{k}(1-\lambda)^k\lambda^{\ell-k}}{\mu(1-\lambda)^2} \\
&= \frac{\sum_\ell p_\ell \sum_{k\leq \ell} k(k-1)\binom{\ell}{k}(1-\lambda)^k\lambda^{\ell-k}}{\mu(1-\lambda)^2} \\
&= \frac{\sum_\ell p_\ell \ell\ (\ell-1)}{\mu} = \nu. \qquad \text{(A.1)}
\end{aligned}$$

To find the diameter in the case $d_{\min} = 1$, we also need to show that $\tilde{q}_1 = \lambda_*$:

$$\begin{aligned}
\tilde{q}_1 = \frac{2\tilde{p}_2}{\tilde{\mu}} &= \frac{2\sum_{\ell\geq 2} p_\ell \binom{\ell}{2}(1-\lambda)^2\lambda^{\ell-2}}{\mu(1-\lambda)^2} \\
&= \frac{1}{\mu}G_p''(\lambda) = G_q'(\lambda) = \lambda_*. \qquad \text{(A.2)}
\end{aligned}$$

We will also need the following relaxation of the notion of 2-core. Let $G = (V, E)$ be a graph. For a given subset $W \subseteq V$, define the W-augmented 2-core to be the maximal induced subgraph of G such that every vertex in $V \setminus W$ has degree at least two, i.e., the vertices in W are not required to verify the minimum degree condition in the definition of the 2-core. The W-augmented 2-core of a graph G will be denoted by $\tilde{G}(W)$.

It is easy to see that the W-augmented 2-core of a random graph $G_n \sim G^*(n, (d_i)_1^n)$, denoted by $\tilde{G}_n(W)$, can be found in the same way as the 2-core, except that now the termination condition is that every nodes outside of W must have degree at least 2, since the half-edges adjacent to a vertex in W are exempt from this restriction. The conditional uniformity property thus evidently holds in this case as well, i.e., for any subset $W \subset V$, the W-augmented 2-core is uniformly random conditional on the W-augmented 2-core half-edge set. We will need the following basic result, the proof of which is easy and can be found for example in [16, Lemma A.7].

Lemma A.2. *Consider a random graph $G_n \sim G(n, (d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.1. For any subset $W \subset V(G_n)$, and any $w \in W$, there exists $C > 0$ (sufficiently large) so that we have*

$$\mathbb{P}\left(e(\tilde{G}_n(W)) - e(\tilde{G}_n(W \setminus \{w\})) \leq C\log n\right) = 1 - o(n^{-1}).$$

Note that the above lemma implies (by removing one vertex from W at a time) that if $|W| = o(n/\log n)$, then w.h.p. the two graphs $\tilde{G}_n$ and $\tilde{G}_n(W)$ have the same degree distribution asymptotic.

APPENDIX B: THE RANDOM GRAPHS $G(N, P)$ AND $G(N, M)$

We derive the results for $G(n, p)$ and $G(n, m)$ from our results for $G(n, (d_i)_1^n)$ by conditioning on the degree sequence. Indeed, we can be more general and consider a random graph G_n with n vertices labeled $[1, n]$ and some random distribution of the edges such that any two graphs on $[1, n]$ with the same degree sequence have the same probability of being attained by G_n. Equivalently, conditioned on the degree sequence, G_n is a random graph with that degree sequence of the type $G(n, (d_i)_1^n)$ introduced in the Introduction. We may thus construct G_n by first picking a random sequence $(d_i)_1^n$ with the right distribution, and then choosing a random graph $G(n, (d_i)_1^n)$ for this $(d_i)_1^n$.

We assume that Condition 1.1 holds in probability:

CONDITION B.1. For each n, let $\mathbf{d}^{(n)} = (d_i^{(n)})_1^n$ be the random sequence of vertex degrees of G_n and $u_k^{(n)}$ be the random number of vertices with degree k. Then, for some probability distribution $(p_r)_{r=0}^{\infty}$ over integers independent of n and with finite mean $\mu := \sum_{k\ge 0} kp_k \in (0, \infty)$, the following holds:

(i) $u_k^{(n)}/n \xrightarrow{p} p_k$ for every $k \ge 1$ as $n \to \infty$;
(ii) For some $\epsilon > 0$, $\sum_{k=1}^{\infty} k^{2+\epsilon} u_k^{(n)} = O_p(n)$;

We first show that for $G(n, p)$ and $G(n, m)$, with $np \to \mu \in (0, \infty)$ and $2m/n \to \mu$, Condition B.1 holds with (p_k) a Poisson distribution with parameter μ, i.e., $p_k = e^{-\mu}\frac{\mu^k}{k!}$. Indeed the fact that Condition B.1(i) holds with such (p_k) follows by elementary estimates of mean and variance done in Example 6.35 of [24] or Theorem 3.1 in [11]. Showing that Condition B.1(ii) holds can be done by similar arguments. Consider $G(n, p)$ (a similar argument holds for $G(n, m)$), we have for all $k \ge 0$ and for n sufficiently large

$$n^{-1}\mathbb{E}u_k^{(n)} = \binom{n-1}{k} p^k (1-p)^{n-1-k} < (\mu+1)^k/k!.$$

Thus, $n^{-1}\sum_{k=1}^{\infty} k^{2+\epsilon}\mathbb{E}u_k^{(n)} = O(1)$, and Condition B.1(ii) holds.

The following lemma is similar to Lemma 8.2 in [22].

LEMMA B.2. *If Condition B.1 holds, we may, by replacing the random graph G_n by other random graphs G_n' with the same distribution, assume* *that the random graphs are defined on a common probability space and that Condition 1.1 holds a.s.*

PROOF. If only Condition 1.1(i) was required, this lemma would be a direct consequence of the Skorohod coupling theorem (Theorem 3.30 [25]) for the random sequence $(u_k^{(n)})_{k=1}^\infty$ in the space $\mathbb{R}_+^\infty$. We now explain how to incorporate Conditions 1.1(ii). Condition B.1 implies that it is possible to find an increasing sequence C_j for $j \geq 1$ diverging to infinity so that considering the sets:

$$A_j = \left\{ (x_k)_{k=1}^\infty \in \mathbb{R}_+^\infty,\ \sum_{k=1}^\infty x_k < \infty,\ \sum_{k=1}^\infty k^{2+\epsilon} x_k \leq C_j \sum_{k=1}^\infty x_k \right\},$$

we have for all n, $\mathbb{P}\left((u_k^{(n)}) \in A_j\right) \geq 1 - (2j)^{-1}$ (note that $\sum_{k=1}^\infty u_k^{(n)} = n$). Let $q_j^{(n)} = \mathbb{P}\left((u_k^{(n)}) \in A_j\right)$ so that $q_{j+1}^{(n)} \geq q_j^{(n)} \geq 1-(2j)^{-1}$ for all $j \geq 1$. For each ℓ, we define an associated finite sequence: $j_i^{(n)}(\ell)$ for $i = 1, \dots, k^{(n)}(\ell)$ such that $j_1^{(n)}(\ell) = 1$ and for $i \geq 1$, $j_{i+1}^{(n)}(\ell) = \min\{j \geq j_i^{(n)}(\ell),\ q_j^{(n)} - q^{(n)}_{j_i^{(n)}(\ell)} \geq \frac{1}{2\ell}\}$ if $q^{(n)}_{j_i^{(n)}(\ell)} < 1 - (2\ell)^{-1}$ and if $q^{(n)}_{j_i^{(n)}(\ell)} \geq 1 - (2\ell)^{-1}$, we set $k^{(n)}(\ell) = i$. Let $\mathcal{J}^{(n)}(\ell) = \{j_1^{(n)}(\ell) = 1, j_2^{(n)}(\ell), \dots, j^{(n)}_{k^{(n)}(\ell)}(\ell)\}$. Note that, since $q_\ell^{(n)} \geq 1 - (2\ell)^{-1}$, we have $k^{(n)}(\ell) \leq \ell$.

We now explicitly construct a 'Skorokhod coupling'. Let θ be a uniform random variable in $[0,1]$ and define the random variable $J^{(n)}(\ell)$ by: $J^{(n)}(\ell) = \min\{j \in \mathcal{J}^{(n)}(\ell), \theta \leq q_j^{(n)}\}$ if $\theta \leq q^{(n)}_{k^{(n)}(\ell)}$ and if $\theta > q^{(n)}_{k^{(n)}(\ell)}$, we set $J^{(n)}(\ell) = \infty$. We set $j_0^{(n)}(\ell) = 0$, $j_i^{(n)}(\ell) = \infty$ for $i > k^{(n)}(\ell)$, $A_0 = \emptyset$ and $A_\infty = \mathbb{R}_+$. With these definitions, we have for all n and $i \geq 1$, $\mathbb{P}(J^{(n)}(\ell) = j_i^{(n)}(\ell)) = \mathbb{P}\left((u_k^{(n)}) \in A_{j_i^{(n)}(\ell)} \backslash A_{j_{i-1}^{(n)}(\ell)}\right)$.

For a given ℓ and for any $i \geq 1$, we define the random variables $\tilde{u}^{(n)}(i) = (\tilde{u}_k^{(n)}(i))_{k \in \mathbb{N}} \in \mathbb{R}_+^\infty$ having the law of $(u_k^{(n)})$ conditioned on the event $\{(u_k^{(n)}) \in A_{j_i^{(n)}(\ell)} \backslash A_{j_{i-1}^{(n)}(\ell)}\}$. Note in particular that by construction, if $i \leq k^{(n)}(\ell)$, we have $\mathbb{P}\left((u_k^{(n)}) \in A_{j_i^{(n)}(\ell)} \backslash A_{j_{i-1}^{(n)}(\ell)}\right) \geq (2\ell)^{-1}$, hence if there exist an infinite sequence of n such that $i \leq k^{(n)}(\ell)$, then we can apply the Skorohod coupling theorem and assume that, along this subsequence, Condition 1.1(i) holds.

We can now combine this coupling with the following one: given θ taken uniformly at random in $[0,1]$, take $\ell = \lceil \frac{1}{2(1-\theta)} \rceil$ and consider $\tilde{u}^{(n)}(J^{(n)}(\ell))$

which has the same law as the original $u^{(n)}$. By construction, Condition 1.1(i) holds. Moreover, we have by construction $J^{(n)}(\ell) \le \ell$ since $q_\ell^{(n)} \ge 1-(2\ell)^{-1} > \theta$, so that $\tilde{u}^{(n)}(J^{(n)}(\ell)) \in A_\ell$ and Condition 1.1(ii) holds.

□

Acknowledgements. This paper is a part of the first author's PhD thesis [1]. He is grateful to Remco van der Hofstad and Laurent Massoulié who reviewed the thesis, and made many valuable comments which helped improve the presentation of this paper. We also thank Omid Amini for helpful comments and discussions. M.L. acknowledges the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-11-JS02-005-01 (GAP project).

REFERENCES

[1] H. Amini. *Epidemics and percolation in random networks.* PhD thesis, ENS-INRIA, Available at *http://www.di.ens.fr/ amini/Publication/Thesis.pdf*, 2011.

[2] H. Amini, M. Draief, and M. Lelarge. Flooding in weighted sparse random graphs. *SIAM Journal on Discrete Mathematics*, 27(1):1–26, 2013.

[3] H. Amini and M. Lelarge. Upper deviations for split times of branching processes. *Journal of Applied Probability*, 49(4):1134–1143, 2012.

[4] E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory. Series A*, 24(3):296–307, 1978.

[5] S. Bhamidi. First passage percolation on locally tree-like networks. I. dense random graphs. *Journal of Mathematical Physics*, 49(12):125218, 2008.

[6] S. Bhamidi and G. H. R. van der Hofstad. First passage percolation on the Erdös-Rényi random graph. *Combinatorics, Probability & Computing*, 20(5):683–707, 2011.

[7] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. Extreme value theory, Poisson-Dirichlet distributions and first passage percolation on random networks. *Advances in applied probability*, 42(3):706–738, 2010.

[8] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. First passage percolation on random graphs with finite mean degrees. *Annals of Applied probability*, 20(5):1907–1965, 2010.

[9] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. Universality for first passage percolation on sparse random graphs. *arXiv preprint arXiv:1210.6839*, 2012.

[10] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[11] B. Bollobás. *Random Graphs.* Cambridge University Press, 2001.

[12] B. Bollobás and W. F. de la Vega. The diameter of random regular graphs. *Combinatorica*, 2(2):125–134, 1982.

[13] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31:3–122, 2007.

[14] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Mathematics*, 1(1):91–113, 2003.

[15] J. Ding, J. H. Kim, E. Lubetzky, and Y. Peres. Diameters in supercritical random graphs via first passage percolation. *Combinatorics, Probability and Computing*, 19:729–751, 2010.

[16] D. Fernholz and V. Ramachandran. The diameter of sparse random graphs. *Random Structures & Algorithms*, 31(4):482–516, 2007.

[17] G. Grimmett and H. Kesten. First-passage percolation, network flows and electrical resistances. *Probability Theory and Related Fields*, 66:335–366, 1984. 10.1007/BF00533701.

[18] O. Häggström and R. Pemantle. First passage percolation and a model for competing spatial growth. *Journal of Applied Probability*, 35(3):683–692, 1998.

[19] R. van der Hofstad, G. Hooghiemstra, and P. V. Mieghem. Distances in random graphs with finite variance degrees. *Random Structures & Algorithms*, 27(1):76–123, 2005.

[20] S. Janson. One, two and three times log n/n for paths in a complete graph with random weights. *Combinatorics, Probability and Computing*, 8(4):347–361, 1999.

[21] S. Janson and M. Luczak. A simple solution to the k-core problem. *Random Structures & Algorithms*, 30:50–62, 2007.

[22] S. Janson and M. J. Luczak. Asymptotic normality of the k-core in random graphs. *Annals of Applied Probability*, 18:1085, 2008.

[23] S. Janson and M. J. Luczak. A new approach to the giant component problem. *Random Structures & Algorithms*, 34(2):197–216, 2009.

[24] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[25] O. Kallenberg. *Foundations of Modern Probability*. 2nd ed., Springer-Verlag, New York, 2002.

[26] H. Kesten. Aspects of first passage percolation. In *École d'été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 125–264. Springer, Berlin, 1986.

[27] A. Klenke and L. Mattner. Stochastic ordering of classical discrete distributions. *To appear in Advances in Applied probability*, 42(2):392–410, 2010.

[28] M. Lelarge. Diffusion and cascading behavior in random networks. *Games and Economic Behavior*, 75(2):752–775, 2012.

[29] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability and Computing*, 7:295–305, 1998.

[30] O. Riordan and N. Wormald. The diameter of sparse random graphs. *Combinatorics, Probability and Computing*, 19:835–926, 2010.

École Polytechnique Fédérale de Lausanne
Lausanne, Switzerland
E-mail: Hamed.Amini@epfl.ch

DYOGENE, INRIA
École Normale Superiéure
Paris, France
E-mail: Marc.Lelarge@ens.fr